

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:22:14 ; Search time 1898.09 Seconds
(without alignments)
13097.759 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCTTCGGGAGCAGCCGCT.....GGAAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_in: *
5: gb_ov: *
6: gb_pal: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_in: *
20: em_on: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1183	99.6	1188	6	AR012692	AR012692 Sequence
2	1160	97.6	1303	9	AF190167	AF190167 Homo sapi
3	1135.2	95.6	1278	9	BC002442	BC002442 Homo sapi
4	1132	95.3	1260	9	AF282596	AF282596 Homo sapi
5	1130.4	95.2	1244	9	AK027405	AK027405 Homo sapi
6	1130.4	95.2	1281	9	BC010152	BC010152 Homo sapi
7	1130	95.1	1262	9	BC014990	BC014990 Homo sapi
8	1120	94.3	1257	9	BC003025	BC003025 Homo sapi
9	867.4	73.0	1253	10	AF23178	AF23178 Mus muscu
10	851.6	71.7	1565	10	BC003425	BC003425 Mus muscu
11	516.6	43.5	691	6	AX261777	AX261777 Sequence
12	515	43.4	518	6	AX193464	AX193464 Sequence
13	455.2	38.3	457	6	AX339946	AX339946 Sequence
14	368.2	31.0	740	6	AX260945	AX260945 Sequence
15	315.6	26.6	18656	2	AC095312	AC095312 Rattus no
16	299.4	25.2	301	6	AX302706	AX302706 Sequence
17	280.8	23.6	1677	8	AF236372	AF236372 Zea mays
18	269.4	22.7	1267	8	AY059109	AY059109 Arabidops
19	269.4	22.7	1503	8	AY034924	AY034924 Arabidops
20	246.6	20.8	12805	1	AE008609	AE008609 Rickettsi
21	240.2	20.2	201050	1	AL646064	AL646064 Ralstonia
22	227.4	19.1	312430	1	RPX002	AJ235221 Rickettsi
23	195	16.4	86765	9	AC004472	AC004472 Homo sapi
24	193.4	16.3	195102	9	AL353795	AL353795 Human DNA
25	193.4	16.3	198829	2	AL391668	AL391668 Homo sapi
26	185.4	15.6	35064	4	AF163772	AF163772 Leishmani
27	185.4	15.6	37852	8	SPHC1665	AL023554 S. pombe
28	180.6	15.2	253	6	A74641	A74641 Sequence 32
29	159.2	13.4	207	6	A77620	A77620 Sequence 32
30	159.2	13.4	44588	2	AX261093	AX261093 Sequence
31	158.2	13.3	148432	3	AC020529	AC020529 Drosophila
32	158.2	13.3	148432	3	AC004642	AC004642 Drosophila
33	158.2	13.3	154840	3	AC099018	AC099018 Drosophila
34	158.2	13.3	300542	3	AE003462	AE003462 Drosophila
35	156.8	13.2	948	1	AF226512	AF226512 Neisseria
36	156.8	13.2	948	1	AF226541	AF226541 Neisseria
37	156.8	13.2	948	6	AX043933	AX043933 Sequence
38	156.8	13.2	948	6	AX043939	AX043939 Sequence
39	156.8	13.2	331601	1	NMA442491	AL162755 Neisseria
40	155.2	13.1	948	1	AF226511	AF226511 Neisseria
41	155.2	13.1	948	1	AF226515	AF226515 Neisseria
42	155.2	13.1	948	1	AF226524	AF226524 Neisseria
43	155.2	13.1	948	1	AF235154	AF235154 Neisseria
44	155.2	13.1	948	1	AF235155	AF235155 Neisseria
45	155.2	13.1	948	1	AF235156	AF235156 Neisseria

ALIGNMENTS

RESULT 1	AR012692	1188 bp	DNA	linear	PAT 05-DEC-1998
LOCUS	AR012692	Sequence 2 from patent US 5763589.			
DEFINITION	AR012692				
ACCESSION	AR012692				
VERSION	AR012692.1	GI:3971010			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1188)				
AUTHORS	Hillman, J.L. and Goll, S.K.				
TITLE	Human membrane protein				
JOURNAL	Patent: US 5763589-A 2 09-JUN-1998;				
FEATURES	Location/Qualifiers				
source	1..1188				
BASE COUNT	268 a 316 c 362 g 237 t				5 others
ORIGIN	/organism="unknown"				

Query Match 99.6%; Score 1183; DB 6; Length 1188;
Best local Similarity 100.0%; Pred. No. 8.9e-271;

Matches 1188: Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Oy	1	GACCTCTGGAGCAGCAGCCCTCCGCTGCTCTGCTTCCGAGAGTCGCTGCGCGGCTGG	60		
Db	1	GACCTCTGGAGCAGCAGCCCTCCGCTGCTCTGCTTCCGAGAGTCGCTGCGCGGCTGG	60		
Oy	61	GAAATGCTGGCGCGCGCGCGCGCGAGCTGGGCGCTTTTGTGAGGCGCTCTCTACT	120		
Db	61	GAAATGCTGGCGCGCGCGCGCGAGCTGGGCGCTTTTGTGAGGCGCTCTCTACT	120		
Oy	121	GACCTCTGGCGCGCGCTCCGCGCGCGCTCCCTGCTGATGGCCGAAACACGCTGGTACTG	180		
Db	121	GACCTCTGGCGCGCGCTCCGCGCGCGCTCCCTGCTGATGGCCGAAACACGCTGGTACTG	180		
Oy	181	TTTCGTGGCCGAGCAGAGAGCCCTGGGTGTGTGAGCAGATGGCGGATTCACCGGATCTCTG	240		
Db	181	TTTCGTGGCCGAGCAGAGAGCCCTGGGTGTGTGAGCAGATGGCGGATTCACCGGATCTCTG	240		
Oy	241	GAGCCTGCTTTGACATCTCTCATCCCTGTGTAGACCGGATCCGATATGTGAGAGTCTC	300		
Db	241	GAGCCTGCTTTGACATCTCTCATCCCTGTGTAGACCGGATCCGATATGTGAGAGTCTC	300		
Oy	301	AAGGAAATTTGATCATCAGCTGCTGAGAGTCGGCTGTGACTCTGACAAATGTAATCTCTG	360		
Db	301	AAGGAAATTTGATCATCAGCTGCTGAGAGTCGGCTGTGACTCTGACAAATGTAATCTCTG	360		
Oy	361	CAAAATCGATGAGTCTTTTACCTGCGCATCATGACCCCTTACAGAGCAAGTACGGTGTG	420		
Db	361	CAAAATCGATGAGTCTTTTACCTGCGCATCATGACCCCTTACAGAGCAAGTACGGTGTG	420		
Oy	421	GAGGAGCCTGAGTATGCCCTCACCCAGCTAGCTCAAAACACCATGATGATCAAGCTCGGC	480		
Db	421	GAGGAGCCTGAGTATGCCCTCACCCAGCTAGCTCAAAACACCATGATGATCAAGCTCGGC	480		
Oy	481	AAACTCTCTNTGAGCAAAAGTCTTCGCGGAAACGGAGATCCCTCAATGCCAGATTGGATG	540		
Db	481	AAACTCTCTNTGAGCAAAAGTCTTCGCGGAAACGGAGATCCCTCAATGCCAGATTGGATG	540		
Oy	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTCGCTTATGATGATCAAGAT	600		
Db	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTCGCTTATGATGATCAAGAT	600		
Oy	601	ATCCATGTGCCACCCCGGGTGAAGAGTCTATGCAAGTGCAGAGTGCAGAGCGCGG	660		
Db	601	ATCCATGTGCCACCCCGGGTGAAGAGTCTATGCAAGTGCAGAGTGCAGAGCGCGG	660		
Oy	661	AAACGGGACCAAGTCTTCTAGTGTGAGGGAGCCGAGATCGGCCATCAATGTGCGAGAA	720		
Db	661	AAACGGGACCAAGTCTTCTAGTGTGAGGGAGCCGAGATCGGCCATCAATGTGCGAGAA	720		
Oy	721	GGAAGAAACAGAGCCCGAGATCTGCGCTCCGAAGCAGAAAGGCTGAGACGATAAATCAG	780		
Db	721	GGAAGAAACAGAGCCCGAGATCTGCGCTCCGAAGCAGAAAGGCTGAGACGATAAATCAG	780		
Oy	781	GCAGAGAGAGAGCCAGTCTGCTGGGGAAGCCCAAGGCTAAAGCTGAAGCTATTCGA	840		
Db	781	GCAGAGAGAGAGCCAGTCTGCTGGGGAAGCCCAAGGCTAAAGCTGAAGCTATTCGA	840		
Oy	841	ATCTGTGCTGACGCTCTGACACAACTAATGAGATGACAGAGCTTCACTGACTGTGGCC	900		
Db	841	ATCTGTGCTGACGCTCTGACACAACTAATGAGATGACAGAGCTTCACTGACTGTGGCC	900		
Oy	901	GAGCAGTATGTACGCGCTTCTCCAAACTGGCCAGAGACTCCAACTATCTACTGCC	960		
Db	901	GAGCAGTATGTACGCGCTTCTCCAAACTGGCCAGAGACTCCAACTATCTACTGCC	960		
Oy	961	TCCAAACCTGGGAGTGTACACAGCATGTGGCTCAGGCCATGGGTATATGAGAGCCCTC	1020		
Db	961	TCCAAACCTGGGAGTGTACACAGCATGTGGCTCAGGCCATGGGTATATGAGAGCCCTC	1020		
Oy	1021	ACCAAAAGCCCGAGTGCAGAGAGCTCCAGACTCACTCTCCAGTGGAGAGCAGAGATGTC	1080		
Db	1021	ACCAAAAGCCCGAGTGCAGAGAGCTCCAGACTCACTCTCCAGTGGAGAGCAGAGATGTC	1080		

Oy	1081	CAGGCTACAGATGCAAGTNTGATGAGAGAACTTGATCGAGTCAAGATGAGTACTGAGAC	1140		
Db	1081	CAGGCTACAGATGCAAGTNTGATGAGAGAACTTGATCGAGTCAAGATGAGTACTGAGAC	1140		
Oy	1141	TGGGCTTCCCGAGAGTCTGGGAGCAAGAGAGCAGATTTTCCTGATT	1188		
Db	1141	TGGGCTTCCCGAGAGTCTGGGAGCAAGAGAGCAGATTTTCCTGATT	1188		
RESULT 2					
AF190167					
LOCUS					
DEFINITION					
Homo sapiens membrane associated protein SLP-2 (HUSLP2) mRNA,					
complete cds.					
ACCESSION					
AF190167					
VERSION					
AF190167.1 GI:6456117					
KEYWORDS					
SOURCE					
ORGANISM					
human.					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE					
1 (bases 1 to 1303)					
Wang, Y. and Morrow, J.S.					
Identification and characterization of human SLP-2, a novel					
homologue of stomatin (band 7.2b) present in erythrocytes and other					
tissues					
JOURNAL					
MEDLINE					
2 (bases 1 to 1303)					
Wang, Y. and Morrow, J.S.					
Direct Submission					
Submitted (25-SEP-1999) Pathology, Yale Medical School, 310 Cedar					
Street, New Haven, CT 06510, USA					
FEATURES					
source					
1..1303					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/chromosome="9"					
/tissue_type="heart muscle"					
1..1303					
/gene="HUSLP2"					
64..1134					
/gene="HUSLP2"					
/note="stomatin-like protein 2; widely distributed					
peripheral membrane protein; similar to human erythrocyte					
stomatin and MEC 1 of Caenorhabditis elegans; thought to					
be involved in mechanoreception or lipid domain					
organization"					
/codon_start=1					
/product="membrane associated protein SLP-2"					
/protein_id="AAF09142.1"					
/db_xref="GI:6456118"					
/translation="MLRAARGTGALLRGLSLASGRAPRRASGLPRNTVVLFPPOQ					
EAMVVERMRGRFRILIEPGILNILLIPDLIRVYSLKEIVINPEGSATVLDNLTQID					
GVLYLRINDPYKASGYEDPEAYVQLAQTMRSELKSLDKVREPRESINASIVDA					
INQADQWIRLRLRIEIKDHYPPRVKSMQVQAEERKRRTVLESCTRESAINVA					
EGRKQAOILASAEAKAEQINQAAAGASVLAIAKAAEIRILAAALYQHNQDAASL					
TVAEQVSAFSLKLANSDNTILLPSNPGDVTSMVAQMGVYALTKAPVPGTDPDSLSG					
SSRDVQGTDAISDEELDRVKMS"					
217					
misc-feature					
/gene="HUSLP2"					
/note="alternative translation initiation site"					
391					
misc-feature					
/gene="HUSLP2"					
/note="alternative translation initiation site"					
BASE COUNT					
315 a 338 c 378 g 272 t					
ORIGIN					
Query Match					
Best Local Similarity 99.4%; Pred. No. 2; 6e-265;					
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;					

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:23:18 ; Search time 1896.25 seconds
(without alignments)
13110.468 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCTTCTGGGACNACCGCT.....GGAAGCAGATTTTCCTGATT 1188

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length DB	ID	Description
-----	-----	-----	-----	-----

1	1188	100.0	1188	6	AR012692	Sequence	1188 bp	DNA	linear	PAT 05-DEC-1998
2	857	72.1	1257	9	BC003025	Homo sapi				
3	857	72.1	1260	9	AF282596	Homo sapi				
4	857	72.1	1262	9	BC014990	Homo sapi				
5	857	72.1	1278	9	BC002442	Homo sapi				
6	857	72.1	1303	9	AF190167	Homo sapi				
7	806	67.8	1244	9	AK027405	Homo sapi				
8	806	67.8	1281	9	BC010152	Homo sapi				
9	454	38.2	457	6	AX339946	Sequence				
10	365	30.7	518	6	AX193464	Sequence				
11	331	27.9	691	6	AX261777	Sequence				
c 12	250	21.0	301	6	AX302706	Sequence				
13	169	14.2	740	6	AX260945	Sequence				
c 14	149	12.5	55406	2	AC026501	Homo sapi				
c 15	149	12.5	86765	9	AC004472	Homo sapi				
c 16	149	12.5	195102	9	AL353795	Human DNA				
17	149	12.5	198829	2	AL391668	Homo sapi				
18	84	7.1	207	6	AX261093	Sequence				
19	74	6.2	253	6	A74641	Sequence 32				
20	74	6.2	253	6	A77620	Sequence 32				
21	53	4.5	1253	10	AF323178	Mus muscu				
22	53	4.5	1565	10	BC003425	Mus muscu				
23	53	4.5	124230	10	AC005259	Mouse BAC				
24	33	2.8	193972	2	AC094506	Rattus no				
c 25	23	1.9	193972	2	AC094506	Rattus no				
26	21	1.8	28958	6	AR044578	Sequence				
27	21	1.8	28958	6	147768	Sequence 6				
28	21	1.8	28958	6	150958	Sequence 6				
29	21	1.8	28958	6	163356	Sequence 6				
30	21	1.8	28958	6	170387	Sequence 6				
31	21	1.8	28958	6	185639	Sequence 6				
32	21	1.8	28958	6	188045	Sequence 4				
33	21	1.8	28958	6	190322	Sequence 6				
34	21	1.8	49377	6	188042	Sequence 1				
35	21	1.8	67523	1	SCU24241	Sorangium c				
c 36	21	1.8	135005	9	HS860F19	Human DNA				
37	21	1.8	151915	2	AC095033	Homo sapi				
c 38	21	1.8	160852	9	AC020915	Homo sapi				
c 39	21	1.8	163903	9	AC007387	Homo sapi				
40	21	1.8	171964	2	AL451003	Homo sapi				
c 41	21	1.8	179307	2	AC104027	Homo sapi				
42	21	1.8	185232	2	AC104005	Homo sapi				
c 43	21	1.8	186656	2	AC095312	Rattus no				
c 44	21	1.8	214042	2	AC023149	Homo sapi				
c 45	21	1.8	221880	2	AC105706	Rattus no				

ALIGNMENTS

RESULT 1	AR012692	Sequence 2 from patent US 5763589.	1188 bp	DNA	linear	PAT 05-DEC-1998
LOCUS	AR012692	Sequence 2 from patent US 5763589.	1188 bp	DNA	linear	PAT 05-DEC-1998
DEFINITION	AR012692	Sequence 2 from patent US 5763589.	1188 bp	DNA	linear	PAT 05-DEC-1998
ACCESSION	AR012692	Sequence 2 from patent US 5763589.	1188 bp	DNA	linear	PAT 05-DEC-1998
VERSION	AR012692.1	GI:3971010				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1188)					
AUTHORS	Hillman, J.L. and Goli, S.K.					
TITLE	Human membrane protein					
JOURNAL	Patent: US 5763589-A 2 09-JUN-1998;					
FEATURES	Location/Qualifiers					
source	1. 1188					
BASE COUNT	268 a	316 c	362 g	237 t	5 others	
ORIGIN						

Query Match 100.0%; Score 1188; DB 6; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 GGCCTTCTGGGAGCACCCTCCGCTCGCTCTGTTGGTTCCGGAGGTCCGCTCGGCGGTGG 60
Db 1 GGCCTTCTGGGAGCACCCTCCGCTCGCTCTGTTGGTTCCGGAGGTCCGCTCGGCGGTGG 60
Oy 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Oy 121 GGCCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GGCCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Oy 181 TTCGTGCCCGCAGCAGAGGCGCTGGGTGGTGGAGGCAATGGGCGCATTCACCGGATCTCG 240
Db 181 TTCGTGCCCGCAGCAGAGGCGCTGGGTGGTGGAGGCAATGGGCGCATTCACCGGATCTCG 240
Oy 241 GAGCCTGCTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC 300
Db 241 GAGCCTGCTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC 300
Oy 301 AAGGAAATGTTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAAGTCTG 360
Db 301 AAGGAAATGTTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAAGTCTG 360
Oy 361 CAATTCGATGGAGTCTTTACCTCGCATCATGAGTACGACCTTACAAGGCAAGCTACGGTGTG 420
Db 361 CAATTCGATGGAGTCTTTACCTCGCATCATGAGTACGACCTTACAAGGCAAGCTACGGTGTG 420
Oy 421 GAGGACCTGAGTATGCGCTCACCAGCTAGCTCAACAAACCATGAGATCAGAGTCTGGC 480
Db 421 GAGGACCTGAGTATGCGCTCACCAGCTAGCTCAACAAACCATGAGATCAGAGTCTGGC 480
Oy 481 AAATCTCTNTGGCAAAAGTCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGGTAT 540
Db 481 AAATCTCTNTGGCAAAAGTCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGGTAT 540
Oy 541 GCCATCAACCAAGCTGTGACTGTGGGTATCGCGTGCCTNCGTTATGAGATCAAGAT 600
Db 541 GCCATCAACCAAGCTGTGACTGTGGGTATCGCGTGCCTNCGTTATGAGATCAAGAT 600
Oy 601 ATCCATGTGCCACCGCGGTGAAGAGTCTATGAGATGCAAGTGGAGGAGCAGCGGCGG 660
Db 601 ATCCATGTGCCACCGCGGTGAAGAGTCTATGAGATGCAAGTGGAGGAGCAGCGGCGG 660
Oy 661 AAACGGCCACAGTCTAGAGTCTGAGGCGACCCGAGAGTCCGCCATCAATGTGCGACAA 720
Db 661 AAACGGCCACAGTCTAGAGTCTGAGGCGACCCGAGAGTCCGCCATCAATGTGCGACAA 720
Oy 721 GGGAGAAACAGGCCAGATCCTTGGCCTCCGAGCAGAAAAAGGCTGAACAGATAAATCAG 780
Db 721 GGGAGAAACAGGCCAGATCCTTGGCCTCCGAGCAGAAAAAGGCTGAACAGATAAATCAG 780
Oy 781 GCAGCAGAGAGGCGCAGTGTGCTGCGGAAGGCCAAGGCTAAGCTGAAGCTATTTCGA 840
Db 781 GCAGCAGAGAGGCGCAGTGTGCTGCGGAAGGCCAAGGCTAAGCTGAAGCTATTTCGA 840
Oy 841 ATCTCTGGCTGACGTCTGACACACATATATGAGATGCGAGCTTCACTGACTGTGGCC 900
Db 841 ATCTCTGGCTGACGTCTGACACACATATATGAGATGCGAGCTTCACTGACTGTGGCC 900
Oy 901 GAGCAGTATGTACGGCGTCTTCCAAATCGCCAGGACTCCAACTATCTACTGCGCC 960
Db 901 GAGCAGTATGTACGGCGTCTTCCAAATCGCCAGGACTCCAACTATCTACTGCGCC 960
Oy 961 TCCAACTCTGGCGATGTACACAGATGCTGCGCTCAGGCCATGGGTGTATATGGAGCCCTC 1020
Db 961 TCCAACTCTGGCGATGTACACAGATGCTGCGCTCAGGCCATGGGTGTATATGGAGCCCTC 1020
Oy 1021 ACCAAAGCCCGAGTCCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTC 1080
Db 1021 ACCAAAGCCCGAGTCCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTC 1080
```

```
Oy 1081 CAGGTACAGATGCAAGTNTTGTATGAGGAAGTCTGATCGAGTCAAGATGAGTGTAGTGGAGC 1140
Db 1081 CAGGTACAGATGCAAGTNTTGTATGAGGAAGTCTGATCGAGTCAAGATGAGTGTAGTGGAGC 1140
Oy 1141 TGGCTTNGCCAGGAGTCTGGGACAGGAAGCAGATTTTCTCTGATT 1188
Db 1141 TGGCTTNGCCAGGAGTCTGGGACAGGAAGCAGATTTTCTCTGATT 1188
```

RESULT 2

BC003025 1257 bp mRNA linear PRI 12-JUL-2001

LOCUS Homo sapiens, stomachin-like 2, clone MGC:4191 IMAGE:2821269, mRNA,

DEFINITION complete cds.

ACCESSION BC003025

VERSION BC003025.1 GI:12804332

KEYWORDS MGC.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1257)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/BTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott zuiderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 9 Row: 1 Column: 7

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9652258.

FEATURES source

Location/Qualifiers

1..1257

/organism="Homo sapiens"

/db_xref="LocusID:30968"

/db_xref="taxon:9606"

/clone="MGC:4191 IMAGE:2821269"

/tissue_type="Lung, small cell carcinoma"

/clone_lib="NIH_MGC_7"

/lab_host="DH10B-R"

/note="vector: pOTB7"

23..1093

/codon_start=1

/product="stomatin-like 2"

/protein_id="AAH03025.1"

/db_xref="GI:12804333"

/translation="MLAARAAGTGALLRGLSLASGRAPRRASGGLPRNTVVLFPVPO

CDS

/translation="MLAARAAGTGALLRGLSLASGRAPRRASGGLPRNTVVLFPVPO

EAWVVRMGFRHRIEPLNLPVLDRIYVOSLKEIVINPVQSVATIDNVTLD

GLVLRIMDPYKASYGVDEPYAVTQLAOTMRSELKLSLDKVFRESINASIVDA

INAAACDWGIRCLRYETKDIHVPVKESQMOMVEAPRRKRTVLESEGTRESAINVA

EGKQKQAILASEAKEAQINQAAGEASVLAkakAKAEATRIILAAALTOINGDAASL

TVAEQYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQMGVYGTAKPVPPTDSSLSC

Query Match 2.0%; Score 7; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 AKAKAKA 255
 Db 2 AKAKAKA 8

```

RESULTS
US-08-346-849-25
; Sequence 25, Application US/08346849
; Patent No: 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,849
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; 346-849-25
  
```


SSRDVOGTDASLDEELDRVKMS"									
BASE COUNT	309 a	324 c	364 g	260 t					
ORIGIN									
Query Match 72.1%; Score 857; DB 9; Length 1257;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	140	GCCGCGCTCTCTGGATTGCCCGAACCCTGGTACTGTTCGTGCGCAGCAGGAGG	199						
Db	99	GCCGCGCTCTCTGGATTGCCCGAACCCTGGTACTGTTCGTGCGCAGCAGGAGG	158						
QY	200	CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTCGGAGCCCTGGTTTGAACATCC	259						
Db	159	CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTCGGAGCCCTGGTTTGAACATCC	218						
QY	260	TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTCATCAACG	319						
Db	219	TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTCATCAACG	278						
QY	320	TGCGTAGCAGTCGGCTGTGACTCTCGCAATGTAACTCTGCAATCGATGGAGTCCTTT	379						
Db	279	TGCGTAGCAGTCGGCTGTGACTCTCGCAATGTAACTCTGCAATCGATGGAGTCCTTT	338						
QY	380	ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG	439						
Db	339	ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG	398						
QY	440	TCACCCAGCTAGCTCAACAAACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG	499						
Db	399	TCACCCAGCTAGCTCAACAAACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG	458						
QY	500	TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	559						
Db	459	TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	518						
QY	560	ACTGCTGGGGTATCCGCTGCCCTNCGTTATGAGATCAAGGATATCCATGTGCCACCCGGG	619						
Db	519	ACTGCTGGGGTATCCGCTGCCCTNCGTTATGAGATCAAGGATATCCATGTGCCACCCGGG	578						
QY	620	TGAAGAGTCTATGCAGATGCAGTGCAGGCGAGCGCGCAACAGGCGCCACAGTTCTAG	679						
Db	579	TGAAGAGTCTATGCAGATGCAGTGCAGGCGAGCGCGCAACAGGCGCCACAGTTCTAG	638						
QY	680	AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCAGA	739						
Db	639	AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCAGA	698						
QY	740	TCCTGGCCTCCGAAGCAGAAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGAGTG	799						
Db	699	TCCTGGCCTCCGACGAGAAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGAGTG	758						
QY	800	CAGTTCCTGGCGAAGGCCAAGGCTAAAGCTTGAAGCTATTTCGAATCCTGGCTCAGCTCTGA	859						
Db	759	CAGTTCCTGGCGAAGGCCAAGGCTAAAGCTTGAAGCTATTTCGAATCCTGGCTCAGCTCTGA	818						
QY	860	CACAACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGCAGCAGTATGTCAAGCGGT	919						
Db	819	CACAACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGCAGCAGTATGTCAAGCGGT	878						
QY	920	TCTCCAACCTGGCCAGGACTTCCAACTATCTACTTCCCTCCAAACCTGGCGATGTCA	979						
Db	879	TCTCCAACCTGGCCAGGACTTCCAACTATCTACTTCCCTCCAAACCTGGCGATGTCA	938						
QY	980	CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTTCAACAAAGCCCGATGCCAG	1039						
Db	939	CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTTCAACAAAGCCCGATGCCAG	998						
QY	1040	GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCAAGT	1098						
Db	999	GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCAAGT	1057						

RESULT	3										
AF282596											
LOCUS											
DEFINITION	Homo sapiens stomatin-like protein 2 (SLP2) mRNA, linear										
ACCESSION	AF282596										
VERSION	AF282596.1										
KEYWORDS	GI:9652258										
SOURCE	human.										
ORGANISM	Homo sapiens										
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.											
REFERENCE	1 (bases 1 to 1260)										
AUTHORS	Owczarek,C.N., Treutlein,H.R., Portbury,K.J., Gulluyan,L.M., Kola,I. and Hertzog,P.J.										
TITLE	A novel member of the STOMATIN/EPB72/mec-2 family, stomatin-like 2 (STOML2), is ubiquitously expressed and localizes to HSA chromosome 9p13.1										
JOURNAL	Cytogenet. Cell Genet. 92 (3-4), 196-203 (2001)										
MEDLINE	21328876										
PUBMED	11435687										
REFERENCE	2 (bases 1 to 1260)										
AUTHORS	Owczarek,C.M., Treutlein,H.R., Portbury,K.J., Gulluyan,L.M., Kola,I. and Hertzog,P.J.										
TITLE	Direct Submission										
JOURNAL	Submitted (26-JUN-2000) Monash Institute of Reproduction and Development, Monash University, Monash Medical Centre, 246 Clayton Road, Clayton, Victoria 3168, Australia										
FEATURES	Location/Qualifiers										
source	1..1260										
/organism="Homo sapiens"											
/db_xref="taxon:9606"											
/chromosome="9"											
/map="9p13.1"											
/tissue_type="brain"											
/dev_stage="fetus"											
gene	1..1260										
/gene="SLP2"											
CDS	35..1105										
/gene="SLP2"											
/note="member of the stomatin/band7/EPB72/MEC2 family; similar to the predicted Caenorhabditis elegans protein F30A10.5, to Caenorhabditis elegans MEC-2, and to the Homo sapiens stomatin protein"											
/codon_start=1											
/product="stomatin-like protein 2"											
/protein_id="AAF91466.1"											
/db_xref="GI:9652259"											
/translation="MLARAAGTGALLRLGSLASGRAPRRASSGLPRNTVVLFPVQQ EAWVERMGRFRILEPLNLILPVLDIRIVQSLKEIVINVPQSAVTLDNVTLOID GVLRLMDPYKASYGVDEPEYAVTQLAQTMRSLEGLKSLDKVFRERESINATVDA INQAADCWGRICLRYEIKDTHVPPRVKSEMQMVEARRKRTVLESEGTRESAINVA EGKKAOILASEAEKAEQINOAGEASAVLAKAKAKAEAIRILAAALTOHNGDAASL TVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAPVPGTPDSLSG SSRDVQGTDAASLDEELDRVKMS"											
1222..1227											
/gene="SLP2"											
polya_signal											
BASE COUNT	303 a	325 c	367 g	265 t							
ORIGIN											
Query Match 72.1%; Score 857; DB 9; Length 1260;											
Best Local Similarity 99.8%; Pred. No. 0;											
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;											
QY	140	GCCGCGCTCTCTGGATTGCCCGAACCCTGGTACTGTTCGTGCGCAGCAGGAGG	199								
Db	111	GCCGCGCTCTCTGGATTGCCCGAACCCTGGTACTGTTCGTGCGCAGCAGGAGG	170								
QY	200	CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTCGGAGCCCTGGTTTGAACATCC	259								
Db	171	CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTCGGAGCCCTGGTTTGAACATCC	230								

Qy	260	TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATTTGTCATCAACG	319
Db	231	TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATTTGTCATCAACG	290
Qy	320	TGCTGAGCAGTCCGCTGTGACTCTCACAATGTAACTCTGCAATCGATGGAGTCCCTTT	379
Db	291	TGCTGAGCAGTCCGCTGTGACTCTCACAATGTAACTCTGCAATCGATGGAGTCCCTTT	350
Qy	380	ACCTGCGCATCATGACCCCTTACAAGCAACAGTACGGTGTGGAGCCCTTGAGTATCCG	439
Db	351	ACCTGCGCATCATGACCCCTTACAAGCAACAGTACGGTGTGGAGCCCTTGAGTATCCG	410
Qy	440	TCACCCAGCTAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTTTGGACAAG	499
Db	411	TCACCCAGCTAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTTGACAAG	470
Qy	500	TCTTCCGGGAACGGAGTCCCTGAAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	559
Db	471	TCTTCCGGGAACGGAGTCCCTGAAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	530
Qy	560	ACTGCTGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCCGG	619
Db	531	ACTGCTGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCCGG	590
Qy	620	TGAAAGAGTCTATCAGATGACAGTGCAGGCGAGCGCGGAAACGGGCCACAGTTCTAG	679
Db	591	TGAAAGAGTCTATCAGATGACAGTGCAGGCGAGCGCGGAAACGGGCCACAGTTCTAG	650
Qy	680	AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAACAGGCCCCAGA	739
Db	651	AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAACAGGCCCCAGA	710
Qy	740	TCCTGGCTCCGAAGCAAGAGCTTCAACAGATAAATCAGGCAGCAGGAGGCCAGTG	799
Db	711	TCCTGGCTCCGAAGCAAGAGCTTCAACAGATAAATCAGGCAGCAGGAGGCCAGTG	770
Qy	800	CAGTTCTGGGCAAGCCAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGCAGCTCTGA	859
Db	771	CAGTTCTGGGCAAGCCAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGCAGCTCTGA	830
Qy	860	CACAACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGAGCAGTATGTACGCGGT	919
Db	831	CACAACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGAGCAGTATGTACGCGGT	890
Qy	920	TCTCCAACTGGCCAGGAGCTCCAACTACTCTCTACTGCTCCCTCCAACTGGCGATGCA	979
Db	891	TCTCCAACTGGCCAGGAGCTCCAACTACTCTCTACTGCTCCCTCCAACTGGCGATGCA	950
Qy	980	CCAGCATGGTGGCTCAGGCCATGGGTATATGGAGCCCTCACCAAGCCCGCAGTGCAG	1039
Db	951	CCAGCATGGTGGCTCAGGCCATGGGTATATGGAGCCCTCACCAAGCCCGCAGTGCAG	1010
Qy	1040	GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGATGTCCAGGTCACAGATCAAGT	1098
Db	1011	GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGATGTCCAGGTCACAGATCAAGT	1069
RESULT	4		
LOCUS	BC014990	1262 bp	mRNA
DEFINITION	Homo sapiens, stomatin-like 2, clone MGC:23225	IMAGE:4907745	linear PRI 29-OCT-2001
ACCESSION	BC014990		complete cds.
VERSION	BC014990.1	GI:15929069	
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1262)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (01-OCT-2001)		National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgabs-r@mail.nih.gov
	Tissue Procurement: ATCC
	cDNA Library Preparation: Rubin Laboratory
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
	info@cgsc.bc.ca
	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Series: IRAL Plate: 34 Row: p Column: 12	
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.	
FEATURES	Location/Qualifiers
source	1..1262
	/organism="Homo sapiens"
	/db_xref="LocusID:30968"
	/db_xref="taxon:9606"
	/clone="MGC:23225 IMAGE:4907745"
	/tissue_type="Pancreas, epithelioid carcinoma"
	/clone_lib="NIH_MGC_42"
	/lab_host="DH10B-R"
	/note="Vector: pOTB7"
	33..1103
	/codon_start=1
	/product="stomatin-like 2"
	/protein_id="AAH14990.1"
	/db_xref="GI:15329070"
	/translation="MLAARAAGTGALLRGLSLASGRAPRRASGGLPRNTVVLFPVPOO EAWVERMGRFHLRLEPLVDRIYVQSLKEIVNPQESAVTLDNVTLOID GLVLRIMDPKASYGVDEPYAVTOLAOTMRSELCKSLDKVFRRESGLNLSAIVDA INQAADCWGIRCLRYEIKDIHPVRKESQMOMOVEARRKRVLESEGTRESAINVA EKKQAQILASEAEKEQINQAGEASAVLAKAKAEAIRILAAALTQHNQDAAASL TVAEQYVSFAFSKLAKDSNTLLPSNFGDVTSMVAQMGVYGTALTAPVPGTSPDSLSSG SSRDVQGTDSLDDELDRVKMS"
CDS	305 a 326 c 367 g 264 t
BASE COUNT	
ORIGIN	
Query Match	72.1%; Score 857; DB 9; Length 1262;
Best Local Similarity	99.8%; Pred. No. 0;
Matches	957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	140 GCCGCGCTCTCTGGATTGCCCGCAACACCGTGGTACTTCTGCTGCCGAGCAGGAGG 199
Db	109 GCCGCGCTCTCTGGATTGCCCGCAACACCGTGGTACTTCTGCTGCCGAGCAGGAGG 168
QY	200 CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 259
Db	169 CCTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 228
QY	260 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGATCTCAAGGAAATTTGTCATCAACG 319
Db	229 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGATCTCAAGGAAATTTGTCATCAACG 288
QY	320 TGCTGAGCAGTCCGCTGTGACTCTCGACAAATGTAACTCTGCAAAATCGATGGAGTCCCTTT 379
Db	289 TGCTGAGCAGTCCGCTGTGACTCTCGACAAATGTAACTCTGCAAAATCGATGGAGTCCCTTT 348

; TITLE: N/A
 ; PUBLICATION INFORMATION:
 ; DOCUMENT NUMBER: PCT/US92/08069
 ; FILING DATE: 14-AUG-1993
 ; US-08-152-488-3

Query Match 2.0%; Score 7; DB 1; Length 29;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
 Db 4 AKAKAKA 10

RESULT 7
 US-08-152-488-3
 ; Sequence 3, Application US/08152488
 ; Patent No. 5534619
 ; GENERAL INFORMATION:
 ; APPLICANT: Wakefield, Thomas W.
 ; APPLICANT: Andrews, Philip C.
 ; APPLICANT: Stanley, James C.
 ; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
 ; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
 ; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Benita J. Rohm, Esq.
 ; STREET: 512 Springfield Avenue
 ; CITY: Cranford
 ; STATE: New Jersey
 ; COUNTRY: United States of America
 ; ZIP: 07016-1811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 6; ASCII (DOS)Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/152,488
 ; FILING DATE: 12-NOV-1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/08069
 ; FILING DATE: 14-AUG-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rohm, Benita J.
 ; REGISTRATION NUMBER: 28,664
 ; REFERENCE/DOCKET NUMBER: RM-7WG
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 908-276-3344
 ; TELEFAX: 908-276-5543
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 29 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: N/A
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: N/A
 ; PUBLICATION INFORMATION:
 ; AUTHORS: N/A


```
QY 380 ACCTGGCATCATGACCCCTTACAAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCGG 439
Db 349 ACCTGGCATCATGACCCCTTACAAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCGG 408
QY 440 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
Db 409 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 468
QY 500 TCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATCCCATCAACCAAGTGTGTG 559
Db 469 TCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATCCCATCAACCAAGTGTGTG 528
QY 560 ACTGCTGGGGTATCCGCTGCGCTTCATGATGATCAAGATATCCATGTCGCCACCCCGGG 619
Db 529 ACTGCTGGGGTATCCGCTGCGCTTCATGATGATCAAGATATCCATGTCGCCACCCCGGG 588
QY 620 TCAAAAGTCTATGAGATGAGGAGTGGAGGAGAGCGGGGAAAGCGGCGACAGTCTTAG 679
Db 589 TCAAAAGTCTATGAGATGAGGAGTGGAGGAGAGCGGGGAAAGCGGCGACAGTCTTAG 648
QY 680 AGTCTGAGGGGACCGAGAGTGGCGCATCAATGTGGCAGAGGAAGAAACAGGCGCCAGA 739
Db 649 AGTCTGAGGGGACCGAGAGTGGCGCATCAATGTGGCAGAGGAAGAAACAGGCGCCAGA 708
QY 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGCAGAGAGAGGCGCAGTG 799
Db 709 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGCAGAGAGAGGCGCAGTG 768
QY 800 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGAGCTCTCA 859
Db 769 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGAGCTCTCA 828
QY 860 CACAACATAATGGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTACGCGGT 919
Db 829 CACAACATAATGGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTACGCGGT 888
QY 920 TCTCCAACTGCCAAGAGCTCAACACTATCTACTGCTCCCTCAACCTCGCGGATGTCA 979
Db 889 TCTCCAACTGCCAAGAGCTCAACACTATCTACTGCTCCCTCAACCTCGCGGATGTCA 948
QY 980 CCAGATGTGCTGCTCAGGCGCATGGGTATATGAGGCGCTCACCAGAGCCCGAGTCCAG 1039
Db 949 CCAGATGTGCTGCTCAGGCGCATGGGTATATGAGGCGCTCACCAGAGCCCGAGTCCAG 1008
QY 1040 GGACTCCAGACTCACTCCAGTGGGAGCAGCAGAGATGTCAGGAGTACAGATGCAAGT 1098
Db 1009 GGACTCCAGACTCACTCCAGTGGGAGCAGCAGAGATGTCAGGAGTACAGATGCAAGT 1067

RESULT 5
BC002442
LOCUS Homo sapiens, stomatin-like 2, clone MGC:1179 IMAGE:3346384, mRNA,
DEFINITION complete cds.
ACCESSION BC002442
VERSION BC002442.1 GI:12803254
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1278)
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J.J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, B.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 5 Row: m Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9652258.

FEATURES
Location/Qualifiers
1..1278
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:1179 IMAGE:3346384"
/tissue_type="skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
47..1117
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH02442.1"
/db_xref="GI:12803255"
/translation="MLAARAGTCALLRGLSLASGRAPRRASSGLPNTVVLFPQQ
EAWVVRGFRHLEPGLNLIPLVDRIIVQSLKEIVINVPESAVTLDNVTIQLD
GVLYLRINDPYKASYVEDPEYATQLAOTMSRLESLDKVRESRETSALNDA
INQADGWIRCLRYEIKDIHVPRKESQMQVEARRKATVLESRETSALNDA
EGKQAOILASEAEKAEQINOAGSAVLAARAKAEAIRILAAALTOHNGDAASL
TVAEQVVSFASKLAKDSNTILLPSNPGDVTNVAQAMGVYALTRAPVPGTPTDLSLSS
SSRDVQGTDSALDELDVRKMS"
BASE COUNT 309 a 330 c 373 g 266 t
ORIGIN

Query Match 72.1%; Score 857; DB 9; Length 1278;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCTCTCTGGATTGCCCGAAACACCGTGTGTCTGTCGCGCAGCAGGAG 199
Db 123 GCCGCGCTCTCTGGATTGCCCGAAACACCGTGTGTCTGTCGCGCAGCAGGAG 182
QY 200 CCTGGGTGTGAGCGAATGGCGGATTCACCGGATCTCTGGAGCCTGTGTAACATCC 259
Db 183 CCTGGGTGTGAGCGAATGGCGGATTCACCGGATCTCTGGAGCCTGTGTAACATCC 242
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGATCTCAAGGAATGTGCATCAACG 319
Db 243 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGATCTCAAGGAATGTGCATCAACG 302
QY 320 TGCCTGAGCAGTCGGCTGTGACTCTCGAACAATGTACTTGCAGAAATCGATGGAGTCTTT 379
Db 303 TGCCTGAGCAGTCGGCTGTGACTCTCGAACAATGTACTTGCAGAAATCGATGGAGTCTTT 362
QY 380 ACTCGGCATCATGGACCTTACAAGGCAAGCTACGGTGTGGAGGACCTGTGATGCGCG 439
Db 363 ACTCGGCATCATGGACCTTACAAGGCAAGCTACGGTGTGGAGGACCTGTGATGCGCG 422
QY 440 TCACCCAGCTACTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
Db 423 TCACCCAGCTACTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTCTGGACAAG 482

```
Qy 500 TCTTCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
Db 483 TCTTCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 542
Qy 560 ACTGCTGGGGTATCCGCTGCCCTTTCATGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 543 ACTGCTGGGGTATCCGCTGCCCTTTCATGATCAAGGATATCCATGTGCCACCCCGGG 602
Qy 620 TGAAGAGTCTATGTCAGATGCAGGTGGAGGAGCGCGGGAAGCGGCCACAGTTCCTAG 679
Db 603 TGAAGAGTCTATGTCAGATGCAGGTGGAGGAGCGCGGGAAGCGGCCACAGTTCCTAG 662
Qy 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGGAAGAACAGGCCCCAGA 739
Db 663 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGGAAGAACAGGCCCCAGA 722
Qy 740 TCCTGGCCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGAGAGGCCAGTG 799
Db 723 TCCTGGCCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGAGAGGCCAGTG 782
Qy 800 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA 859
Db 783 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA 842
Qy 860 CACAACATAATGAGATGCAGCAGCTTCACTGACGTGTGGCCGAGCAGTATGTACGCGGT 919
Db 843 CACAACATAATGAGATGCAGCAGCTTCACTGACGTGTGGCCGAGCAGTATGTACGCGGT 902
Qy 920 TCTCCAACTGCCAAGACATCCACATATCCTACTGCCCTCCAACTCCAGCTGATGCA 979
Db 903 TCTCCAACTGCCAAGACATCCACATATCCTACTGCCCTCCAACTCCAGCTGATGCA 962
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATGAGGCGCTCACCAAGGCCCGAGTGCAG 1039
Db 963 CCAGCATGGTGGCTCAGGCCATGGGTGTATGAGGCGCTCACCAAGGCCCGAGTGCAG 1022
Qy 1040 GGACTCCAGACTCACTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1023 GGACTCCAGACTCACTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1081

RESULT 6
AF190167 AF190167 1303 bp mRNA linear PRI 13-MAR-2000
DEFINITION Homo sapiens membrane associated protein SLP-2 (HUSLP2) mRNA,
complete cds.
ACCESSION AF190167
VERSION AF190167.1 GI:6456117
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1303)
Wang, Y. and Morrow, J.S.
Identification and characterization of human SLP-2, a novel
homologue of stomatin (band 7.2b) present in erythrocytes and other
tissues
J. Biol. Chem. 275 (11), 8062-8071 (2000)
20179914
MEDLINE
J. Biol. Chem. 275 (11), 8062-8071 (2000)
REFERENCE
2 (bases 1 to 1303)
Wang, Y. and Morrow, J.S.
Direct Submission
TITLE Submitted (25-Sep-1999) Pathology, Yale Medical School, 310-Cedar
Street, New Haven, CT 06510, USA
JOURNAL Location/Qualifiers
1..1303
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/tissue_type="heart muscle"
1..1303
/gene="HUSLP2"

gene
```

```
CDS
64..1134
/gene="HUSLP2"
/note="stomatin-like protein 2; widely distributed
peripheral membrane protein; similar to human erythrocyte
stomatin and MEC 1 of Caenorhabditis elegans; thought to
be involved in mechanoreception or lipid domain
organization"
/codon_start=1
/product="membrane associated protein SLP-2"
/protein_id="AAF09142.1"
/db_xref="GI:6456118"
/translation="MLAARAGTGCALLRSLGSLASGRAPRRASSGLPRNTVVLFVPOQ
EAWVERMGRFHRILEPGLNLIIPVLDRIYVQSLKEIVINPQSAVTLDNVTLQID
GVLYLRIMDPYKASGYVEDPEYATQLAQTMRSELKLSLDKVFRRERESINASTIDA
INOADCWGIRCLRYETKDIHVPRVKESMOMOVEABRRKRATVLESEGTRESAINVA
EGKQAOILASEAEKAEQINOAGEASAVLAKAKAEAIRLLAAALTOHNGDAASL
TVABQYVSATFSKSLAKDSNTILLPSNPGDVTSWAQAMGVGYALTAKVPVPTDLSLSSG
SSRDVQGTDAASLDELDLRVXMS"
217
/misc_feature
/gene="HUSLP2"
/note="alternative translation initiation site"
391
/misc_feature
/gene="HUSLP2"
/note="alternative translation initiation site"
315 a 338 c 378 g 272 t
BASE COUNT
ORIGIN
Query Match 72.1%; Score 857; DB 9; Length 1303;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 140 GCCGGCCCTCTCTGGATTGCCCGCAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 199
Db 140 GCCGGCCCTCTCTGGATTGCCCGCAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 199
Qy 200 CCTGGGTGGTGAGCGAATGGCCGATTCACCCGGATCCTGGAGCCTGTTTGAACATCC 259
Db 200 CCTGGGTGGTGAGCGAATGGCCGATTCACCCGGATCCTGGAGCCTGTTTGAACATCC 259
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATGTGTCAACG 319
Db 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATGTGTCAACG 319
Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCMAATCGATGAGTCCCTT 379
Db 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCMAATCGATGAGTCCCTT 379
Qy 380 ACCTGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 380 ACCTGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGACTCGGCAAACTCTCTTGGACAAG 499
Db 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGACTCGGCAAACTCTCTTGGACAAG 499
Qy 500 TCTTCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGTGTGCT 559
Db 500 TCTTCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGTGTGCT 559
Qy 560 ACTGCTGGGGTATCCGCTGCCCTTTCATGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 560 ACTGCTGGGGTATCCGCTGCCCTTTCATGATCAAGGATATCCATGTGCCACCCCGGG 619
Qy 620 TGAAGAGTCTATGTCAGATGCAGGTGGAGGAGCGCGGGAAGCGGCCACAGTTCCTAG 679
Db 620 TGAAGAGTCTATGTCAGATGCAGGTGGAGGAGCGCGGGAAGCGGCCACAGTTCCTAG 679
Qy 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGGAAGAACAGGCCCCAGA 739
Db 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGGAAGAACAGGCCCCAGA 739
Qy 740 TCCTGGCCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGAGAGGCCAGTG 799
```

```
Db 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATAAATCAGCAGCAGGAGGCGCAGTG 799
Qy 800 CAGTTCTGGCAGCCAGGCTAAAGCTGAAGCTATTTCGAATCTCGCTGCGAGCTCTGA 859
Db 800 CAGTTCTGGCAGCCAGGCTAAAGCTGAAGCTATTTCGAATCTCGCTGCGAGCTCTGA 859
Qy 860 CACAACATAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Db 860 CACAACATAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Qy 920 TCTCCAAACTGCCAAGACATCCACACATATCTACTGCTCCCTCCACCTGCGGATGCA 979
Db 920 TCTCCAAACTGCCAAGACATCCACACATATCTACTGCTCCCTCCACCTGCGGATGCA 979
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAAGCCCGAGTCCAG 1039
Db 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAAGCCCGAGTCCAG 1039
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGTTACAGATGCAAGT 1098
Db 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGTTACAGATGCAAGT 1098

RESULT 7
AK027405 1244 bp mRNA linear PRI 15-MAY-2001
LOCUS Homo sapiens cDNA FLJ14499 fis, clone NT2RM1000080, weakly similar
DEFINITION to UNC-1 PROTEIN.
ACCESSION AK027405
VERSION AK027405.1 GI:14042059
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA,
clone_lib:NT2RM1 clone:NT2RM1000080.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1244)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
source 1..1244
organism="Homo sapiens"
db_xref="taxon:9606"
clone="NT2RM1000080"
cell_line="NT2"
clone_type="teratocarcinoma"
clone_lib="NT2RM1"
note="cloning vector: pUC19FL3-mRNA from uninduced NT2
neuronal precursor cells."
35..1105
note="unnamed protein product"
/codon_start=1
```

```
/protein_id="BAB55091.1"
/db_xref="gi:14042060"
/translation="MLAARAGTGALLIRGSLASGRAPRRASSGLPRNTVFLVFPQO
EAWVERMRGHRILEPLGLNLIIPVLDRIYVQSLKEIVINVPQSAVTLDNVTLOID
GLVLRIMDPYKASYGVGDEPEYATOLAOITMRSELKLSLDKVPRESLNASLIDA
INQAAOCWGIRCLREYKDIHVPVRKESMOMQVEARRKRTATVLESGTRESAINVA
EGKKAQILASEAEKEDINQAGEASAVLAKAKAKAEAIRLLAALTOHNGDAASL
TVABQYVSFAFKLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAPYPGTSDLSLS
SSRDVQGTGTDASLDELDRVMS"
BASE COUNT 284 a 327 c 368 g 265 t
ORIGIN
Query Match 67.8%; Score 806; DB 9; Length 1244;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 956; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 140 GCCGGCCTCTCTCGATTGCCCGAAGACCGTGTACTGTTCGTGCGCAGCAGGAGG 199
Db 111 GCCGGCCTCTCTCGATTGCCCGAAGACCGTGTACTGTTCGTGCGCAGCAGGAGG 170
Qy 200 CCTGGTGTGTGAGCGAATGGGCCGATTCCACCGGATCTCGAGGCTGCTTTGAACATCC 259
Db 171 CCTGGTGTGTGAGCGAATGGGCCGATTCCACCGGATCTCGAGGCTGCTTTGAACATCC 230
Qy 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 319
Db 231 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 290
Qy 320 TGCCTGAGCAGTCCGGCTGTGACTCTCGACATGTACTCTGCAATTCGATGAGTCTTT 379
Db 291 TGCCTGAGCAGTCCGGCTGTGACTCTCGACATGTACTCTGCAATTCGATGAGTCTTT 350
Qy 380 ACTCGGCATCATGACCCCTTACAAGGCAAGTACGGTGTGGAGGACCTGAGTATGCCG 439
Db 351 ACTCGGCATCATGACCCCTTACAAGGCAAGTACGGTGTGGAGGACCTGAGTATGCCG 410
Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
Db 411 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 470
Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGTATGTGGATGCCATCAACCAAGCTGTG 559
Db 471 TCTTCCGGGAACGGAGTCCCTGAATGCCAGTATGTGGATGCCATCAACCAAGCTGTG 530
Qy 560 ACTGTGGGGTATCCGCTGCCCTTCATGAGATCAAGGATATCCATGTGCCACCCCGG 619
Db 531 ACTGTGGGGTATCCGCTGCCCTTCATGAGATCAAGGATATCCATGTGCCACCCCGG 590
Qy 620 TGAAGAGTCTATGAGATGTCAGGTGGAGGAGCGGCGGAAACGGCCACAGTTCTTAG 679
Db 591 TGAAGAGTCTATGAGATGTCAGGTGGAGGAGCGGCGGAAACGGCCACAGTTCTTAG 650
Qy 680 AGTCTGAGGGACCGAGAGTCCCATCAATGTGGCAGAAAGGAGAAACAGGCCACAG 739
Db 651 AGTCTGAGGGACCGAGAGTCCCATCAATGTGGCAGAAAGGAGAAACAGGCCACAG 710
Qy 740 TCCTGCGCTCCCAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGGAGGCGCAGTG 799
Db 711 TCCTGCGCTCCCAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGGAGGCGCAGTG 770
Qy 800 CAGTTCTGCGGAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCTCGCTGCGAGCTCTGA 859
Db 771 CAGTTCTGCGGAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCTCGCTGCGAGCTCTGA 830
Qy 860 CACAACATAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Db 831 CACAACATAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 890
Qy 920 TCTCCAAACTGCCAAGGACTCCACACATATCTACTGCTCCCTCCACCTGCGGATGCA 979
Db 891 TCTCCAAACTGCCAAGGACTCCACACATATCTACTGCTCCCTCCACCTGCGGATGCA 950
```


RESULT 9
AX339946
LOCUS AX339946 457 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 193 from Patent WO0196388.
ACCESSION AX339946
VERSION AX339946.1 GI:18135927
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196388-A 193 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..457
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 123 c 137 g 73 t 2 others
ORIGIN
Query Match 38.2%; Score 454; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 2e-262;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 633 GCAGATGCGAGTGGAGGAGCGGGGAAACGGCCACAGTTCTAGAGTCTGAGGGGAC 692
Db 1 GCAGATGCGAGTGGAGGAGCGGGGAAACGGCCACAGTTCTAGAGTCTGAGGGGAC 60
QY 693 CCGAGAGTCGGCCATCAATGTGGCAGAGGAGGAAACAGAGCCAGATCCTGGCCTCCGA 752
Db 61 CCGAGAGTCGGCCATCAATGTGGCAGAGGAGGAAACAGAGCCAGATCCTGGCCTCCGA 120
QY 753 AGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGGAGGAGGAGTCTTGGCGAA 812
Db 121 AGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGGAGGAGTCTTGGCGAA 180
QY 813 GGCCAAAGGCTAAAGTGAAGCTATTGGAATCTGCGTGTGAGTGTGACACACATAATGG 872
Db 181 GGCCAAAGGCTAAAGTGAAGCTATTGGAATCTGCGTGTGAGTGTGACACACATAATGG 240
QY 873 AGATGCGAGCTTCACTGACTGTGGCCAGCAGTATGTGAGGCGGTTCTCCAAACTGGC 932
Db 241 AGATGCGAGCTTCACTGACTGTGGCCAGCAGTATGTGAGGCGGTTCTCCAAACTGGC 300
QY 933 CAAGGACTCCAACTATCTTACTGTGCGCTTCCAACTGCGGATGTGAGGAGTGTGGC 992
Db 301 CAAGGACTCCAACTATCTTACTGTGCGCTTCCAACTGCGGATGTGAGGAGTGTGGC 360
QY 993 TCAGGCGATGGGTGTATATGAGGAGCTTCAACAAAGGCGGAGTGTGAGGAGTGTGAG 1052
Db 361 TCAGGCGATGGGTGTATATGAGGAGCTTCAACAAAGGCGGAGTGTGAGGAGTGTGAG 420
QY 1053 ACTCTCCAGTGGGAGCAGCAGAGATCTCCAGGT 1086
Db 421 ACTCTCCAGTGGGAGCAGCAGAGATCTCCAGGT 454
RESULT 10
AX193464
LOCUS AX193464 518 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 1031 from Patent WO0149716.
ACCESSION AX193464
VERSION AX193464.1 GI:15211415
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher,M.J., Xu,J. and King,G.E.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 1428 04-OCT-2001;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 518)
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 1031 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..518
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 123 a 136 c 149 g 109 t 1 others
ORIGIN
Query Match 30.7%; Score 365; DB 6; Length 518;
Best Local Similarity 99.4%; Pred. No. 1.5e-208;
Matches 515; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 200 CTGGGTGTGGAGCGAATGGCCGATTCACCGGATCTCGAGCTGTGGAGCTGTGGAATCC 259
Db 1 CTGGGTGTGGAGCGAATGGCCGATTCACCGGATCTCGAGCTGTGGAGCTGTGGAATCC 60
QY 260 TCATCCCTCTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 319
Db 61 TCATCCCTCTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 120
QY 320 TGCTGAGCAGTGGGTGTGACTCTCGACATGTAACTGTGCAATTCGATGGAGTCTTT 379
Db 121 TGCTGAGCAGTGGGTGTGACTCTCGACATGTAACTGTGCAATTCGATGGAGTCTTT 180
QY 380 ACTCTGGCATATGGACCTTACAAGCAAGCTACCGTGTGGAGGAGCTGTGATGATGCGG 439
Db 181 ACTCTGGCATATGGACCTTACAAGCAAGCTACCGTGTGGAGGAGCTGTGATGATGCGG 240
QY 440 TCACCCAGCTAGTCTCAACCAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
Db 241 TCACCCAGCTAGTCTCAACCAACCATGAGATCAGAGCTCGGCAAACTCTCTCTGACAAAG 300
QY 500 TCTTCGGGAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
Db 301 TCTTCGGGAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 360
QY 560 ACTCTGGGTATCCCTGCTGCTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
Db 361 ACTCTGGGTATCCCTGCTGCTTATGAGATCAAGATATCCATGTGCCACCCCGG 420
QY 620 TGAAGAGTCTATGAGATGAGGTGGAGCGAGCGGGAACGGCCACAGTTCTAG 679
Db 421 TGAAGAGTCTATGAGATGAGGTGGAGCGAGCGGGAACGGCCACAGTTCTAG 480
QY 680 AGTCTGAGGGGAGCGAGAGTGGCCATCAATGTGCA 717
Db 481 AGTCTGAGGGGAGCGAGAGTGGCCATCAATGTGCA 518
RESULT 11
AX261777
LOCUS AX261777 691 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1428 from Patent WO0173027.
ACCESSION AX261777
VERSION AX261777.1 GI:16510744
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher,M.J., Xu,J. and King,G.E.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 1428 04-OCT-2001;

```

CORIXA CORPORATION (US)
Location/Qualifiers
1. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 183 a 168 c 184 g 140 t 16 others
ORIGIN

Query Match 27.9%; Score 331; DB 6; Length 691;
Best Local Similarity 99.7%; Pred. No. 5.5e-188;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 200 CTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 259
Db 1 CTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 60

Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTCGAGAGTCTCAAGGAATTTGATCAACG 319
Db 61 TCATCCCTGTGTAGACCGGATCCGATATGTCGAGAGTCTCAAGGAATTTGATCAACG 120

Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACCTCTGCAATCGATGGAGTCCTTT 379
Db 121 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACCTCTGCAATCGATGGAGTCCTTT 180

Qy 380 ACCTCGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTCGATATGCCG 439
Db 181 ACCTCGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTCGATATGCCG 240

Qy 440 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAGCTCTCTTGGACAAG 499
Db 241 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAGCTCTCTTGGACAAG 300

Qy 500 TCTTCCGGGAACGGAGTCCCTGTAATGCCAGCATGTGGATGCCATCAACCAAGCTGCTG 559
Db 301 TCTTCCGGGAACGGAGTCCCTGTAATGCCAGCATGTGGATGCCATCAACCAAGCTGCTG 360

Qy 560 ACTGCTGGGGTATCCGCTGCCT 581
Db 361 ACTGCTGGGGTATCCGCTGCCT 382

RESULT 12
AX302706/c
LOCUS AX302706 301 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 26 from Patent WO0179286.
ACCESSION AX302706
VERSION AX302706.1 GI:17383208
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and
Repler,W.F.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: WO 0179286-A 26 25-OCT-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .301
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 48 a 96 c 72 g 85 t
ORIGIN

Query Match 21.0%; Score 250; DB 6; Length 301;
Best Local Similarity 99.7%; Pred. No. 5.6e-139;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 626 AGTCTATGAGATGAGGTGGAGGAGCGCGCGGAACGGCCACAGCTTCTAGAGTCTG 685
```

```

Db 301 AGTCTATGAGATGAGGTGGAGGAGCGCGCGGAACGGCCACAGCTTCTAGAGTCTG 242
Qy 686 AGGGACCCGAGAGTCGCCCATCAATGTGGCAGAAGGAAGAAACAGGCCAGATCCTGG 745
Db 241 AGGGACCCGAGAGTCGCCCATCAATGTGGCAGAAGGAAGAAACAGGCCAGATCCTGG 182
Qy 746 CCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGACAGGAGAGGCCAGTGCAGATTC 805
Db 181 CCTCCGAAGCAGAAAGGCTGCACAGATAAATCAGGACAGGAGAGGCCAGTGCAGATTC 122
Qy 806 TGGCGAAGCCCAAGGCTAAAGCTTAAGCTATTCGAATCCTGGCTCCAGCTCTGACACAC 865
Db 121 TGGCGAAGCCCAAGGCTAAAGCTTAAGCTATTCGAATCCTGGCTCCAGCTCTGACACAC 62
Qy 866 ATAATGGAGATGCAGCAGCTTTCACCTGACTGTGGCGAGCAGTATGTACGCGGTTCTCCA 925
Db 61 ATAATGGAGATGCAGCAGCTTTCACCTGACTGTGGCGAGCAGTATGTACGCGGTTCTCCA 2
Qy 926 A 926
Db 1 A 1

RESULT 13
AX260945
LOCUS AX260945 740 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 596 from Patent WO0173027.
ACCESSION AX260945
VERSION AX260945.1 GI:16509912
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher,M.J., Xu,J. and King,G.E.
TITLE Compositions and methods for therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 596 04-OCT-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1. 740
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 189 a 181 c 169 g 159 t 42 others
ORIGIN

Query Match 14.2%; Score 169; DB 6; Length 740;
Best Local Similarity 99.5%; Pred. No. 5.8e-90;
Matches 219; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 247 GGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAA 306
Db 120 GGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAA 179
Qy 307 ATTGTATCAACAGTGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCTGCAAAATC 366
Db 180 ATTGTATCAACAGTGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCTGCAAAATC 239
Qy 367 GATGAGTCTTTACTTGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGAC 426
Db 240 GATGAGTCTTTACTTGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGAC 299
Qy 427 CCTGAGTATGCCGTCAACCCAGCTAGCTCAACCAACCATGA 466
Db 300 CCTGAGTATGCCGTCAACCCAGCTAGCTCAACCAACCATGA 339

RESULT 14
AC026501/c
LOCUS AC026501 55406 bp DNA linear HTG 22-MAR-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-266M12 map 1, LOW-PASS
```



```

SEQUENCE SAMPLING.
AC026501
VERSION AC026501.1 GI:7283191
KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 55406)
JOURNAL Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE Homo sapiens chromosome 1, clone RP11-266M12
2 (bases 1 to 55406)
* 5476 5575: gap of 100 bp
* 5576 6263: contig of 688 bp in length
* 6264 6363: gap of 100 bp
* 6364 7060: contig of 697 bp in length
* 7061 7160: gap of 100 bp
* 7161 7869: contig of 709 bp in length
* 7870 7969: gap of 100 bp
* 7970 8683: contig of 714 bp in length
* 8684 8783: gap of 100 bp
* 8784 9466: contig of 683 bp in length
* 9467 9566: gap of 100 bp
* 9567 10263: contig of 697 bp in length
* 10264 10363: gap of 100 bp
* 10364 11053: contig of 690 bp in length
* 11054 11153: gap of 100 bp
* 11154 11848: contig of 695 bp in length
* 11849 11948: gap of 100 bp
* 11949 12639: contig of 691 bp in length
* 12640 12739: gap of 100 bp
* 12740 13444: contig of 705 bp in length
* 13445 13544: gap of 100 bp
* 13545 14237: contig of 693 bp in length
* 14238 14337: gap of 100 bp
* 14338 15047: contig of 710 bp in length
* 15048 15147: gap of 100 bp
* 15148 15720: contig of 573 bp in length
* 15721 15820: gap of 100 bp
* 15821 16512: contig of 692 bp in length
* 16513 16612: gap of 100 bp
* 16613 17316: contig of 704 bp in length
* 17317 17416: gap of 100 bp
* 17417 18118: contig of 702 bp in length
* 18119 18218: gap of 100 bp
* 18219 18904: contig of 686 bp in length
* 18905 19004: gap of 100 bp
* 19005 19696: contig of 692 bp in length
* 19697 19796: gap of 100 bp
* 19797 20486: contig of 690 bp in length
* 20487 20586: gap of 100 bp
* 20587 21293: contig of 707 bp in length
* 21294 21393: gap of 100 bp
* 21394 22098: contig of 705 bp in length
* 22099 22198: gap of 100 bp
* 22199 22875: contig of 677 bp in length
* 22876 22975: gap of 100 bp
* 22976 23682: contig of 707 bp in length
* 23683 23782: gap of 100 bp
* 23783 24474: contig of 692 bp in length
* 24475 24574: gap of 100 bp
* 24575 25280: contig of 706 bp in length
* 25281 25380: gap of 100 bp
* 25381 26083: contig of 703 bp in length
* 26084 26183: gap of 100 bp
* 26184 26874: contig of 691 bp in length
* 26875 26974: gap of 100 bp
* 26975 27679: contig of 705 bp in length
* 27680 27779: gap of 100 bp
* 27780 28457: contig of 678 bp in length
* 28453 28557: gap of 100 bp
* 28558 29255: contig of 698 bp in length
* 29256 29355: gap of 100 bp
* 29356 30053: contig of 698 bp in length
* 30054 30153: gap of 100 bp
* 30154 30847: contig of 694 bp in length
* 30848 30947: gap of 100 bp
* 30948 31656: contig of 709 bp in length
* 31657 31756: gap of 100 bp
* 31757 32459: contig of 703 bp in length
* 32460 32559: gap of 100 bp
* 32560 33270: contig of 711 bp in length
* 33271 33370: gap of 100 bp
* 33371 34072: contig of 702 bp in length
* 34073 34172: gap of 100 bp
*

TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6433
Center clone name: 266_M_12
-----
* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 677 776: contig of 676 bp in length
* 777 776: gap of 100 bp
* 777 1478: contig of 702 bp in length
* 1479 1578: gap of 100 bp
* 1579 2267: contig of 689 bp in length
* 2268 2367: gap of 100 bp
* 2368 3052: contig of 685 bp in length
* 3053 3152: gap of 100 bp
* 3153 3870: contig of 718 bp in length
* 3871 3970: gap of 100 bp
* 3971 4671: contig of 701 bp in length
* 4672 4771: gap of 100 bp
* 4772 5475: contig of 704 bp in length
*

```

* 34173	34836:	contig of 664 bp	in length	
* 34837	34936:	gap of 100 bp		
* 34937	35630:	contig of 694 bp	in length	
* 35631	35730:	gap of 100 bp		
* 35731	36435:	contig of 705 bp	in length	
* 36436	36535:	gap of 100 bp		
* 36536	37220:	contig of 685 bp	in length	
* 37221	37320:	gap of 100 bp		
* 37321	37982:	contig of 662 bp	in length	
* 37983	38082:	gap of 100 bp		
* 38083	38777:	contig of 695 bp	in length	
* 38778	38877:	gap of 100 bp		
* 38878	39560:	contig of 683 bp	in length	
* 39561	39660:	gap of 100 bp		
* 39661	40347:	contig of 687 bp	in length	
* 40348	40447:	gap of 100 bp		
* 40448	41129:	contig of 682 bp	in length	
* 41130	41229:	gap of 100 bp		
* 41230	41927:	contig of 698 bp	in length	
* 41928	42027:	gap of 100 bp		
* 42028	42733:	contig of 706 bp	in length	
* 42734	42833:	gap of 100 bp		
* 42834	43516:	contig of 683 bp	in length	
* 43517	43616:	gap of 100 bp		
* 43617	44335:	contig of 719 bp	in length	
* 44336	44435:	gap of 100 bp		
* 44436	45123:	contig of 688 bp	in length	
* 45124	45223:	gap of 100 bp		
* 45224	45928:	contig of 705 bp	in length	
* 45929	46028:	gap of 100 bp		
* 46029	46720:	contig of 692 bp	in length	
* 46721	46820:	gap of 100 bp		
* 46821	47524:	contig of 704 bp	in length	
* 47525	47624:	gap of 100 bp		
* 47625	48304:	contig of 680 bp	in length	
* 48305	48404:	gap of 100 bp		
* 48405	49107:	contig of 703 bp	in length	
* 49108	49207:	gap of 100 bp		
* 49208	49873:	contig of 666 bp	in length	
* 49874	49973:	gap of 100 bp		
* 49974	50640:	contig of 667 bp	in length	
* 50641	50740:	gap of 100 bp		
* 50741	51431:	contig of 691 bp	in length	
* 51432	51531:	gap of 100 bp		
* 51532	52222:	contig of 691 bp	in length	
* 52223	52322:	gap of 100 bp		
* 52323	53018:	contig of 696 bp	in length	
* 53019	53118:	gap of 100 bp		
* 53119	53822:	contig of 704 bp	in length	
* 53823	53922:	gap of 100 bp		
* 53923	54617:	contig of 695 bp	in length	
* 54618	54717:	gap of 100 bp		
Query Match				
Best Local Similarity 12.5%: Score 149; DB 2: Length 55406;				
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	640	CAGGTGGAGCAGAGCGCGGAAAGCGGCCACAGTCTAGAGTCTGAGGGACCCGAGAG	699	
DB	38593	CAGGTGGAGCAGAGCGCGGAAAGCGGCCACAGTCTAGAGTCTGAGGGACCCGAGAG	38534	
QY	700	TCGGCCATCAATGTGGCAGAGAGGGAAGAAACAGGCCACAGATCTCGGCTCCGAGCAGAA	759	
DB	38533	TCGGCCATCAATGTGGCAGAGAGGGAAGAAACAGGCCACAGATCTCGGCTCCGAGCAGAA	38474	
QY	760	AAGGCTGAACAGATAAATCAGGCAGCAGG	788	
DB	38473	AAGGCTGAACAGATAAATCAGGCAGCAGG	38445	
RESULT	15			
AC004472/c				
LOCUS	AC004472	86765 bp	DNA	linear PRI 15-JUN-2001

DEFINITION Homo sapiens chromosome 9, P1 clone 11659, complete sequence.

ACCESSION AC004472

VERSION AC004472.1 GI:2984582

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 86765)

AUTHORS Lamerdin, J.E., McCready, P.M., Skowronski, E., Adamson, A.W., Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stilwagen, S., Phan, H., Velasco, N., Ganes, J., Danganan, L., Poundstone, P., Christensen, M., Georgescu, A., Avila, J., Liu, S., Atti, C., Andreise, T., Trankheim, M., Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Kronmiller, B., Arellano, A., Montgomery, M., Ow, D., Nolan, M., Trong, S., Kobayashi, A., Olsen, A.O., and Carrano, A.V.

TITLE Sequence analysis of a human P1 clone containing the XRCC9 DNA repair gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 86765)

AUTHORS Lamerdin, J.E.

TITLE Direct Submission

JOURNAL Submitted (25-MAR-1998) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA

FEATURES

source

1..86765

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="9p13"

/clone="P1.11659"

/sex="Male"

/tissue_type="Fibroblast or foreskin"

/lab_host="NS3516"

/note="P1 clone obtained from Genome Systems, Inc."

/rpt_family="MER4_internal"

83..383

/rpt_family="AluY"

/complement(402..626)

/rpt_family="MER4_internal"

/complement(707..1338)

/rpt_family="HERVH"

/complement(1361..1741)

/rpt_family="HERVH"

/complement(1742..3015)

/rpt_family="HERVH"

/complement(3358..4084)

/rpt_family="HERVH"

/complement(4145..4571)

/rpt_family="HERVH"

/complement(4579..5035)

/rpt_family="LTR7"

/complement(5036..5621)

/rpt_family="MER4_internal"

5634..5733

/rpt_family="AluS"

5734..6025

/rpt_family="AluSx"

/complement(6027..6110)

/rpt_family="MER83"

6113..6662

/rpt_family="MER41A"

/complement(6663..6994)

/rpt_family="MER83"

/complement(7000..7557)

/rpt_family="MER4_internal"

/complement(7560..8234)

/rpt_family="MER4A"

8239..9165

/rpt_family="MSTB-internal"

/complement(9215..9483)

Job time: 7130 sec

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1188	100.0	1188	19	AAV28867	Human integral mem	
2	857	72.1	1337	20	AAV04345	Human secreted pro	
3	857	72.1	1398	22	ABA09225	Human SLP-2 homolo	
4	857	72.1	1398	22	AAK52683	Human polynucleoti	
5	857	72.1	9098	22	AAK34953	CDNA encoding nove	
6	855	72.0	1322	20	AAH04402	Human secreted pro	
7	806	67.8	1244	22	AAH13961	Human cDNA sequenc	
8	793	66.8	1416	22	AAK51699	Human polynucleoti	
9	365	30.7	518	22	AAI29477	Colon tumour relat	

[illegible]

```
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0055984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058666.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060865.
PR 02-OCT-1997; 97US-0061059.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX
XX WPI; 1999-080881/07.
DR P-P5DB; AAW78160.
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 1; Page 201-202; 380pp; English.
XX
XX This sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number, and the clone it is derived
CC from, are detailed in the descriptor line. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAX04302) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic
CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 86
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX04311 for described uses).
XX
XX Sequence 1337 BP; 338 A; 346 C; 377 G; 273 T; 3 other;
SQ
```

```
Query Match 72.1%; Score 857; DB 20; Length 1337;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCGGCGCTCTCTGTGATTGCCCGGAAACACCGTGTACTGTTCGGCCGACGAGAGG 199
Db |||||||
QY 152 GCGGCGCTCTCTGTGATTGCCCGGAAACACCGTGTACTGTTCGGCCGACGAGAGG 211
Db |||||||
QY 200 CTGTGGTGTGGAGCAATGGCCGATTCACCGGATCTCGAGCTGTGCTTTCAACATCC 259
Db |||||||
QY 212 CTGTGGTGTGGAGCAATGGCCGATTCACCGGATCTCGAGCTGTGCTTTCAACATCC 271
Db |||||||
QY 260 TCATCCTCTGTGTAGACCGGATTCGAGATGTGAGATGTGAGATGTGATCAACG 319
Db |||||||
QY 272 TCATCCTCTGTGTAGACCGGATTCGAGATGTGAGATGTGAGATGTGATCAACG 331
Db |||||||
QY 320 TGCTGAGCAGTGGGTGTGACTCTCGAATGTAACTGTGCAAAATCGATGGAGTCTTT 379
Db |||||||
QY 332 TGCTGAGCAGTGGGTGTGACTCTCGAATGTAACTGTGCAAAATCGATGGAGTCTTT 391
Db |||||||
QY 380 ACCTGCGCATGAGCCCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 439
Db |||||||
```

```
Db 392 acctgcgcatcatgacacctataaagggaagctcaaggtgtgaggaccctgagatgctg 451
QY 440 TCACCCAGCTAGCTCAACAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
Db |||||||
Db 452 tcaccagctagtctcaacaacacatgagatcagagctcgcaactctctcttgacaaag 511
QY 500 TCTTCGGGAAAGGGAGTCCCTGAATGCGAGCATTTGGATGCCATCAACCAAGCTGCTG 559
Db |||||||
Db 512 tcttcgggaaaggagtcctgaaatgagcagcattgtggtgcatcaacaaagctgctg 571
QY 560 ACTGCTGGGTTATCCGCTCCCTTCGTTATGATCAAGATATCCATGTGCCACCCCGGG 619
Db |||||||
Db 572 actgtggtgtatccgtctccgttatgadatcaagatatccatgtgccccccggg 631
QY 620 TGAAGAGTCTATGAGATGCGAGTGGAGCGAGCGGGGAAACGGCCACAGTCTCTAG 679
Db |||||||
Db 632 tgaagagctctatgagatgcaggtgagagcagcgcggaacaggccacagttctag 691
QY 680 AGTCTGAGGGGCCCGAGAGTGGCCATCAATGTGCGAGGAAGGAAGAACAGGCCGAGA 739
Db |||||||
Db 692 agtctgaggggacccgagagtgccatcaatgtgagaaagggaagaaacagggccaga 751
QY 740 TCCTGGCTCCGGAAGCAGAAAAGGCTGAACAGATAAATCAGGCAGGAGAGGCCAGTG 799
Db |||||||
Db 752 tcttgccctccgaagcagaaaaggctgaacagataaatcaggcagcagagagccagtg 811
QY 800 CAGTTCCTGGCGAGGCGCAAGGCTAAAGCTGAAGCTATTGAATCTCGCTGCGAGCTCTGA 859
Db |||||||
Db 812 cagttctggaaggccaaaggctaaagctgaagctattcgaatcctggtcgagctctga 871
QY 860 CACAACATAATGGAGATGCGAGCTTCACTGACTGTGGCCGAGCAGTATGTGAGGCGCT 919
Db |||||||
Db 872 cacaacataatggagatgcagcagcttctactgactgtggtgagcagcagtagtcagcgt 931
QY 920 TCTCCAACTGGCCAAAGGACTCCACACATATCTACTGCTCCCAACCTGGCGAGTCTCA 979
Db |||||||
Db 932 tctccaaactggcgaaggactccactctactgctccctccacccctggcgagtgca 991
QY 980 CCAGCATGTGTGCTCAGGCGCATGGGTGTATATGAGGCCCTCACCAGCCCGAGTCCAG 1039
Db |||||||
Db 992 ccagcatgtgttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1051
QY 1040 GGACTCCAGCTACTCTCCAGTGGGAGGAGCAGAGATGTCAGGGGTACAGATGCAATG 1098
Db |||||||
Db 1052 ggaactccagactcactctccagtggtgagcagcagagatgtccaggttacagatgca 1110
XX
RESULT 3
ABR09225/c
ID ABR09225 standard; cDNA; 1398 BP.
XX
AC ABA09225;
XX
XX 11-JAN-2002 (first entry)
XX
DE Human SLP-2 homologue-encoding cDNA, SEQ ID NO:1001.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
OS Homo sapiens.
XX
```

PN WO200157188-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US03800.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457740/49.
DR P-PSDB; ABB11981.
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX Claim 1; Page 851; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.

Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;

Query Match 72.1%; Score 857; DB 22; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGGCCCTCTGGATTGCCCGAACAACCGTGGTACTGTTCTGCCGAGCAGGAGG 199
Db 1254 GCCGGCCCTCTGGATTGCCCGAACAACCGTGGTACTGTTCTGCCGAGCAGGAGG 1195
Qy 200 CTGGGTGGTGGAGGAATGGCCGATTCACCGGATCCTGGAGCGTGGTTGAACATCC 259
|||||

Db 1194 CCTGGTGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGTTTGAACATCC 1135
Qy 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATGTTGTCATCAACG 319
Db 1134 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATGTTGTCATCAACG 1075
Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGAGTGGAGTCCCTTT 379
Db 1074 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGAGTGGAGTCCCTTT 1015
Qy 380 ACCTGGCATCATGACCCCTTACAAGGAAGCTACGGTGTGGAGGACCTCTCTNTTGGACAAG 439
Db 1014 ACCTGGCATCATGACCCCTTACAAGGAAGCTACGGTGTGGAGGACCTCTCTNTTGGACAAG 955
Qy 440 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTTGGACAAG 499
Db 954 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTTGGACAAG 895
Qy 500 TCTTCCGGGAACGGGAGTCCCTGTAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
Db 894 TCTTCCGGGAACGGGAGTCCCTGTAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 835
Qy 560 ACTGTGGGGTATCCGCTGCGCTTATGAGATCAAGATATCCATGTGCCACCCCGGG 619
Db 834 ACTGTGGGGTATCCGCTGCGCTTATGAGATCAAGATATCCATGTGCCACCCCGGG 775
Qy 620 TGAAGAGTCTATGATGAGATGAGAGTGGAGGAGCGGCGGAAGCGGCCAGTCTTAG 679
Db 774 TGAAGAGTCTATGATGAGATGAGAGTGGAGGAGCGGCGGAAGCGGCCAGTCTTAG 715
Qy 680 AGTCTGAGGGACCGGAGAGTGGGCTCAATGTGGCAGAAGGAAGAACAGGCCGAGA 739
Db 714 AGTCTGAGGGACCGGAGAGTGGGCTCAATGTGGCAGAAGGAAGAACAGGCCGAGA 655
Qy 740 TCCTGGCTCCGAGCAAGAAAGGCTGAACAGATTAATCAGGCAGCAGGAGGAGGCGAGTG 799
Db 654 TCCTGGCTCCGAGCAAGAAAGGCTGAACAGATTAATCAGGCAGCAGGAGGAGGCGAGTG 595
Qy 800 CAGTTCTGGGAGCGCCAGGCTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGA 859
Db 594 CAGTTCTGGGAGCGCCAGGCTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGA 535
Qy 860 CACAACATAATGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTGTCAGCGGT 919
Db 534 CACAACATAATGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTGTCAGCGGT 475
Qy 920 TCTCCAAACTGCCCAAGAGCTCCCAACTATCTTACTGCCCTCCCAACCTGCGGATGTCA 979
Db 474 TCTCCAAACTGCCCAAGAGCTCCCAACTATCTTACTGCCCTCCCAACCTGCGGATGTCA 415
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGAGCCCGAGTGGCAG 1039
Db 414 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGAGCCCGAGTGGCAG 355
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 354 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 4
AAK52683/c
ID AAK52683 standard; cDNA: 1398 BP.
XX
AC AAK52683;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2212.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.
XX WO200157190-A2.
XX
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US04098.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX DR WPI: 2001-476283/51.
XX DR P-P5DB; AAM79550.
XX
XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX PT useful in diagnosis and gene therapy -
XX
XX PS Claim 1: Page 4572; 6221pp; English.
XX
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX CC (AAM8020) are omitted as the relevant pages from the sequence listing
XX CC were missing at the time of publication.
XX
XX SQ Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;

Query Match 72.1%; Score 857; DB 22; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCGGCGCTCTGTGATGGCCCGGAAACACCGTGTACTGTTCGTGCGCAGCAGGAGG 199
DB 1254 GCGGCGCTCTGTGATGGCCCGGAAACACCGTGTACTGTTCGTGCGCAGCAGGAGG 1195
QY 200 CTGTGGTGTGGAGCGAATGGCGGATTCACCGGATCTCGAGCTGCTTGAACATCC 259
DB 1194 CTGGGTGGAGCGAATGGCGGATTCACCGGATCTCGAGCTGCTTGAACATCC 1135
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAAG 319
DB 1134 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAAG 1075
QY 320 TGCCTGAGCATCGGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGGAGTCTTT 379
DB 1074 TGCCTGAGCATCGGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGGAGTCTTT 1015
QY 380 ACCTGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 439
DB 1014 ACCTGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 955

QY 440 TCACCCAGCTAGCTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTTNTGGACAAAG 499
DB 954 TCACCCAGCTAGCTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTCTGGACAAAG 895
QY 500 TCTTCCGGGAAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
DB 894 TCTTCCGGGAAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 835
QY 560 ACTGTCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGATATCCATGTCACCCCGGG 619
DB 834 ACTGTCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGATATCCATGTCACCCCGGG 775
QY 620 TGAAGAGTCTATGACAGATGACAGTGGAGGAGCGGGGAAACGGGCGACACATGTTCTAG 679
DB 774 TGAAGAGTCTATGACAGATGACAGTGGAGGAGCGGGGAAACGGGCGACACATGTTCTAG 715
QY 680 AGTCTGAGGGGACCGCAGAGTGGCCCATCAATGTGCGAGAGGAAACAGAGCCCGGAG 739
DB 714 AGTCTGAGGGGACCGCAGAGTGGCCCATCAATGTGCGAGAGGAAACAGAGCCCGGAG 655
QY 740 TCTTGGCCTCCCAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGAGAGGCGCAGTG 799
DB 654 TCTTGGCCTCCCAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGAGAGGCGCAGTG 595
QY 800 CAGTTCTGCGAAGCCCAAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGACAGCTTGA 859
DB 594 CAGTTCTGCGAAGCCCAAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGACAGCTTGA 535
QY 860 CACAACATATGAGATGACAGAGCTTCACTGACTGTGCGCGGAGCAGATGTCAGCGCT 919
DB 534 CACAACATATGAGATGACAGAGCTTCACTGACTGTGCGCGGAGCAGATGTCAGCGCT 475
QY 920 TCTTCAAACTGGCAAGGACTCCAACACTATCTACTGCTCCCAACCTCGCGGATGTCA 979
DB 474 TCTTCAAACTGGCAAGGACTCCAACACTATCTACTGCTCCCAACCTCGCGGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAGCCCGGAGCCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAGCCCGGAGCCAG 355
QY 1040 GGACTCCAGACACACACCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 354 GGACTCCAGACACACACCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 5
AAS44953
ID AAS44953 standard; cdna; 9098 BP.
XX AC AAS44953;
XX DT 18-DEC-2001 (first entry)
XX DE cdna encoding novel human secretory protein, Seq ID No 34.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.
XX
OS Homo sapiens.
XX
XX PN WO200166689-A2.
XX PD 13-SEP-2001.
XX PF 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QH, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX WPI: 2001-589934/66.
DR P-PSDB; AAU28053.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Claim 1; SEQ ID No 34; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention.
XX
SQ Sequence 9098 BP; 1918 A; 2538 C; 2590 G; 2052 T; 0 other;

Query Match 72.1%; Score 857; DB 22; Length 9098;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGCGCTCTCTGGATTGCCCGAACCAGCGTGGTACTGTCTGCGCGCAGCAGGAGG 199
Db 151 gcccgcctcctctgattgcccgaaacacgctgttactgtctgctgcgcagcaggagg 210
Qy 200 CTGGGCTGGTGGAGCGATGGCCCATTCACCGGATCTCGGAGCCTGGTTGACATCC 259
Db 211 cctgggtgtgagcgaatggccgattccaccggtatcccgatcctgagcgtgtttgaacatcc 270
Qy 260 TCATCCCTGTGTACACCGGATCCCATATGTGCAGAGTCTCAAGGAATGTGCATCAACG 319

Db 271 tcataccctgtgttagaccggatccgatatgtcgagagttctcaaggaaattgtcatcaacg 330
Qy 320 TGCCTGAGCAGTCGCTGTGACTCTCGACAATGTAACTCTGCAAAATCGAGTCCCTTT 379
Db 331 tgcctgagcagtcgctgtgactctcgcacaaagttaactctgcaaatcgatgagtcctct 390
Qy 380 ACCTGCGCATCATGACCCCTTACAAGGCAAGCTACGGGTGTGGAGGACCCCTGAGTATCGCG 439
Db 391 acctgcgcatcatgaccccttacaaggcaagctacggtgtggaggacctgatgacg 450
Qy 440 TCACCCAGCTAGCTCAACAAACCATGATCAGACTCGGCAAACTCTCTNTGGAGAAAG 499
Db 451 tcaccagtagctcaaacacacatgagatcagagctcggaactctctctggacaag 510
Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTCTGGATGCCATCAACCAAGCTGCTG 559
Db 511 tcttcctgggaacggagtcctcctgaatgccagcatgtgtgagtcacacacacagctgctg 570
Qy 560 ACTGCTGGGGTATCCGCTGCCCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 571 actgctggggtatccgctgcctccgttatgagatcaaggatatccatgtgccaccccg 630
Qy 620 TCAAGAGTCTATGAGATGCGAGTGGAGGCGAGCGGCGGAAACGGCCAGTCTCTAG 679
Db 631 tgaagagttctatgagatgcaggtggaggcagcgcggaacgggcccagttctag 690
Qy 680 AGTCTGAGGGGACCGAGAGTGGCGCATCAATGTGGCAGAAGGAAGAAACAGGCCACGA 739
Db 691 agtctgagggggaccgcagagtcggtccatcaatgtggcagaaagggaagaaacaggcc 750
Qy 740 TCTTGGCTCCGAAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGGAGGCGCAGTG 799
Db 751 tcttgctcccgaaagcagaagaggtgaacagataaaatcaggcagcagagggccagctg 810
Qy 800 CAGTTCTGCGGAAGCGCAGGCTTAAGCTGAAGCTTTCGAATCTGCTGCTGAGCTCTGA 859
Db 811 cagttctggcggaagcgaaggaagctgaagctattcgaatcctggtgagctctga 870
Qy 860 CACAACATTAATGGAGATGCGAGCGTTCAGTACTGTGGCCGAGCAGTATGTCAGCGCT 919
Db 871 cacaacataatggagatgcagcagcttcaactgactgtggcgcagcagtagtgcagcgc 930
Qy 920 TCTCAAACTGCCAAGGACTCCCAACTATCTTACTGCCCTCCAAACCTCGCGGATGTCA 979
Db 931 tctcaaaactggccagagactcccaactatcctactgcctcccaacctcggcagtgta 990
Qy 980 CCAGCATGTGCTCAGGCGCATGGGTGTATATGAGCGCTTACCAAGGCCCGCAGTCCAG 1039
Db 991 ccagcatgggtcagggccatgggtgtatatggagccctcaccacaaagccccagtgcc 1050
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1051 ggactccagactcactctccagtgggagcagcagagatgtccagggtacagatgcaagt 1109

RESULT 6

AAAX04402

ID AAAX04402 standard; DNA; 1322 BP.

XX AC AAAX04402;

XX XX

DT 13-APR-1999 (first entry)

XX XX

DE Human secreted protein gene 35 clone HTXCS21.

XX XX

Human: secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.
OS WO9856804-A1.
PN 17-DEC-1998.
PD 11-JUN-1998; 98WO-US12125.
PF 02-OCT-1997; 97US-0061060.
XX 13-JUN-1997; 97US-0049547.
XX 13-JUN-1997; 97US-0049548.
PR 13-JUN-1997; 97US-0049549.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049606.
PR 13-JUN-1997; 97US-0049607.
PR 13-JUN-1997; 97US-0049608.
PR 13-JUN-1997; 97US-0049609.
PR 13-JUN-1997; 97US-0049610.
PR 13-JUN-1997; 97US-0049611.
PR 13-JUN-1997; 97US-0050566.
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0055984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058668.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060865.
PR 02-OCT-1997; 97US-0061059.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Brewer LA, Ebner R, Ferrrie AM, Feng P, Greene JM, Lafleur DW;
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX WPI: 1999-080881/07.
DR P-PSDB; AAW78217.
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX Claim 1; Page 257-258; 380pp; English.
XX This sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number, and the clone it is derived
CC from, are detailed in the descriptor line. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAX04302) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic
CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 86
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX04311 for described uses).
XX Sequence 1322 BP; 336 A; 337 C; 373 G; 271 T; 5 other;
SQ

Query Match 72.0%; Score 855; DB 20; Length 1322;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 955; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 142 CGCGCTCTCTGGATTGCGCGAAACACCGCTGGTACTGTTCTGTCGCGCAGCAGGAGGCC 201
DB 139 cgcgcctcctctggattgcccgcgaacaccgctggctactgttctgctgcgcagcaggaagcc 198
QY 202 TGGGTGGTGGAGCGGAATGGCGGATTTCCACCGGATTCCTACCGGCTGGTGGTGAACATCTTC 261
DB 199 tgggtggtagcggaatggcggaattccacggatctcctggagcctgtttgaacatcctc 258
QY 262 ATCCCTGTGTAGACCGGATTCGATATGTCAGAGTCTCAAGAAATTTCTCATCAACGTG 321
DB 259 atccctgtgttagaccggatccgatattgtagagagtcctcaagaaatttcatcaacagt 318
QY 322 CCGTACAGCTCGCTGTGACTCTCGACAATGTAACCTCTCAATTCGATCGAGTCTCTTAC 381
DB 319 cctgagcagtcggctgtgactctcgacaattgtaactctgcaaatcgatgagatccctttac 378
QY 382 CTGCGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGACCCCTGAGTATGCCGTG 441
DB 379 ctgcgcacatgagacccttacaaggcaagctacggctgtaggagaccctgagtatgcgctc 438
QY 442 ACCCAGCTAGCTCAACACCATGAGATCAGAGTCTCGGCAAACTCTCTTTGGACAAAGTC 501
DB 439 acccagctagctcaacaacacatgagatcagagctcgcaaaactctctctggacaagtc 498
QY 502 TTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTGAC 561
DB 499 ttccgggaacgggagtcctctgaatgccagcatgtggaatgcacaaacagctgctgac 558
QY 562 TGCTGGGTATCCGCTGCCCTTATGAGATCAAGGATATCCATGTGCCACCCCGGGTG 621
DB 559 tgcctgggtatccgctgcctcgttatgagatcaagatatccatgtgccaccccggtg 618
QY 622 AAGAGCTCTATGAGATGAGTGGAGGAGGAGCGCGGAAACGGGCCACAGTTCTAGAG 681
DB 619 aaagagctctatgagatgagctgtaggagcagagcggcgaacacggccacagttctagag 678
QY 682 TCTGAGGAGCCCGAGAGTCCGCTCAATGTGGCAGAGAGGGAAGAAACAGGCCAGATC 741
DB 679 tctgaggggaccgcagagctcgccatcgaatgtggcagagggggaagaaacagggccagatc 738
QY 742 CTGGCTCCGAGCAGAGAAAGGCTGAACAGATAAATCAGCAGCAGGAGGAGCCAGTGCA 801
DB 739 ctggcctccgagcagaaaaggctgaacagataaatcaggcagcagagagggccagtga 798
QY 802 GTTCTGGCGAAGCCCAAGGCTAAAGCTGAAGCTATTTCGAATCTCTGCTGCGAGTCTTGACA 861
DB 799 gttctggcgaaggccaaggctaaagctgagctattcgaaatcctgctgctgagctgaca 858
QY 862 CAACATAATGGAGATGACGAGCTTCACTGCTGCTGCGCAGCAGTATGTACGCGGCTTC 921
DB 859 caacataatggagatgcagcagcttcactgactgtggcgcagcagtatgtcagcgagttc 918
QY 922 TCCAAACTGCCAAGGACTCCAACTATCTCTACTGCTCCCTCCAACTCCGCGATGTCACC 981
DB 919 tccaaactggcgaaggactccaaactatctactgctccctccaaacccctggcgatgcacc 978
QY 982 AGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGCCCCAGTCCGAGGG 1041
DB 979 agcatggtggctcagggccatgggtgtatgtgagccctcaccagaagccccagtgccagg 1038
QY 1042 ACTCCAGACTCTCTCTCCAGTGGGAGCAGCAGAGATGTCCAGGCTACAGATGCAAGT 1098
DB 1039 actccagactcactctccagtggggagcagcagagatgtccaggggtacagatgcaagt 1095
RESULT 7
AAH13961
ID AAH13961 standard; cDNA; 1244 BP.
XX
AC AAH13961;

XX
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:11009.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 11009; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1244 BP; 284 A; 327 C; 368 G; 265 T; 0 other;
SQ

Query Match 67.8%; Score 806; DB 22; Length 1244;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 956; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 140 GCCGGCTCTCTGGATTGCCCGCAACACCGGTGGTACTTCTGTCGCCGAGCAGGAGG 199
Db 111 gccgcctctctgattgtcccgaaacacccgtgttctgttcgtccgcagcaggagg 170
Qy 200 CTGTGGTGGTGGAGCGAATGGCGGATTCACCGGATCCTGGAGCCTGGCTTTGAACATCC 259
Db 171 cctgggtggtgagcgaatggcgcattccaccggatcctcgagcctggttgaacatcc 230

QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATTTGTATCAACG 319
DB 231 tcatccctgtgttagaccggtatccgatatgtgcagagttctcaaggaaatctgcatcaacg 290
QY 320 TGGCTGAGCAGTCGGCTGTGACTCTCTGACAAATTAACCTGCAAAATCGATGGAGTCTCTT 379
DB 291 tgcctgagcagtcgctgtgactctcgacaataaactctgcaaatcgatggagtccttt 350
QY 380 ACCTGGCATCATGAGACCCCTTACAAGCAAGCTACCGTGTGGAGGACCCCTGATATGCCG 439
DB 351 acctgcgcatcatggacccttacaaggcaagctacggtgtggaggaccctgagtatgcg 410
QY 440 TCACCCAGCTAGCTCAACCAACCATGAGATCAGATCGATCGCAAACTCTCTNTGGACAAG 499
DB 411 tcacccagctagctcaacaacccatgagatcagagctcggcaaacctctcctgcaaaag 470
QY 500 TCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
DB 471 tcttcgggggaacgggagtcctgaatgccagcatgtggatgccatcaaccaagctgctg 530
QY 560 ACTGCTGGGTATCCGCTGCCTNCGTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
DB 531 actgctggggtatccgctgcctcctcgttatgagatcaaggatatccatgtgccacccgg 590
QY 620 TGAAGAGTCTATGCAGATCAGGTGGAGCGAGCGCGAAACGGCCACACAGTTCTAG 679
DB 591 tgaagagtcctatgcagatgcaggtggaggcagagcgcggaacggcgccacagctctag 650
QY 680 AGTCTAGGGAGCCCGGAGAGTCGGCCATCAATGTGGCAGAAGGAGAAACAGAGCCCGAGA 739
DB 651 agtctgaggggaccggagagtcggccatcaatgtggcagaagggaagaaacagggccaga 710
QY 740 TCCTGGCTCCGGAAGCAAGAGCTGAACAGATAAATCAGGACGAGGAGGAGCCAGTG 799
DB 711 tcctggcctccgaagcagaagagctgaacagataaaatcaggcagcagagagccagtg 770
QY 800 CAGTTCCTGGCGAAGGCCAAGCTAAAGCTGAAGCTATTTCGAATCTTGGCTGCAGCTCTGA 859
DB 771 cagttctcggcgagggccaaagctaaagctaaagctaaagctaaagctaaagctaaagctaa 830
QY 860 CACAACATAATGAGATGTCAGAGCTTCACTGACTGTGGCGGAGCAGATATGTCTACGCGGT 919
DB 831 cacaacataatgagatgagatgcagcagcttcaactgactgtggcgcagcagtatgtcagcgc 890
QY 920 TCTCAAACTGGCCCAAGGACTCCAACTATCTTACTGCTCCCAACCTCCGCGATGCA 979
DB 891 tctcaaaactggcgaagagactcccaactatcctactgcoctccaaacctggcgatgca 950
QY 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGAGCCCGAGTGGCAG 1039
DB 951 ccagcatgggtgggtcagggccatgggtgtatatggagccctcaccagagcccgagtgccgg 1010
QY 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGGTACAGATCAAGT 1098
DB 1011 ggactccagactcactctccagtggggagcagcagagatgtccagggtacagatgcaagt 1069

RESULT 8
AAK51699
ID AAK51699 standard; cDNA; 1416 BP.
XX AC AAK51699;
XX DT 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 244.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX OS Homo sapiens.

XX WO200157190-A2.
PN 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0683561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR P-PSDB; AAM78566.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX Claim 1; Page 1135-1136; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX SQ Sequence 1416 BP; 335 A; 364 C; 410 G; 307 T; 0 other;

Query Match 66.8%; Score 793; DB 22; Length 1416;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 893; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGGCTCTCTGGATTGCCCGGAAACACCGTGTACTGTTCGTCGCGCAGCAGGAGG 199
Db 150 gccgcctctctgattgcccgaaacacgctgtactgttcgtgcgcgagcaggagg 209
Qy 200 CTTGGGTGGTGAGCGAATGGCCGATTCACCCGGATCCTCGAGCCTGTTTGAACATCC 259
Db 210 cctgggtggtgagcgaaatggccgattccaccgagctcctgagcctggtttgaacatcc 269
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGTCTCAGGAAATGTTCATCAACG 319
Db 270 tcatccctgtgttagaccggatccgatatgtgcagagtctcaaggaaattgtcatcaacg 329
Qy 320 TGCCTGAGCAGTCGCTGTGACTCTCGCAATGTAACTCTGCAATCGATGAGTCCCTTT 379
Db 330 tgcctgagcagtcgctgtgactctcgcaaatgttaactctgcaaatcgatggagtccttt 389
Qy 380 ACCTCGCATCATGACCCCTTACAGGCAAGCTACCGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 390 acctgcgcatcatgaccttacaaggcaagctacggtgtgaggagacctgagtcgagcgg 449
Qy 440 TCACCCAGCTAGCTCAACACCATGAGATCAGACTCGGCNAACCTCTCTNTGGACAAG 499

Db 450 tcaccagctagctcaaaacaaccatgagatcagactcggaactctctctggcaag 509
Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
Db 510 tcttcggggaacggagtgctccctgaatgccagcatgtgtgagatgccatcaaccaagctgctg 569
Qy 560 ACTGCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGGATATCCATGTCACCCCGGG 619
Db 570 actgctggggtatccgctgctccgcttatgagatcaaggatatccatgtgcccacccggg 629
Qy 620 TCAAGAGTCTATGACAGTGCAGTGGAGGTCAGAGCGGCGGAAACGGGCCACAGTCTAG 679
Db 630 tgaagagctctatgagatgcaggtggagggcagagcgcggaacgggcccagttctag 689
Qy 680 AGTCTGAGGGACCGGAGAGTGGCGCATCAATGTGGCAGAAGGAAGAAACAGGCCACGA 739
Db 690 agtctgaggggagcccgagagtcggccatcaatgtgtggcagaagggaagaaagggcccgag 749
Qy 740 TCCTGGCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGCAGCAGGAGGAGGCGAGTG 799
Db 750 tectggcctccgaagcagaaaagcctgaacagataaatcagcagcagcagagggccagtg 809
Qy 800 CAGTTCTGCGGAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCTGCTGCTGAGCTCTGA 859
Db 810 cagttctggcgaagccaaaggctaaagctgaagctatttcgaatcctctgctgagctctga 869
Qy 860 CACAACATAATGGAGATGCAGCAGCTTCACCTGCTGCTGCGCAGCAGTATGTGAGGCGCT 919
Db 870 cacaacataatggagatgcagcagcttcactgactgtgtggcagcagcagtatgtcagcgct 929
Qy 920 TCTCCAAACTGCGCAAGGACTCCAACTATCTACTGCTCCCTCCAACTGCGCTGCGATGCA 979
Db 930 tctccaaactgccaaggaactccaaactatctactgctcccaacctggcgatgtca 989
Qy 980 CCAGCATGCTGCTCAGGCCATGGGTGTATATGAGGAGCCCTCACCAGGCCCCCAGT 1034
Db 990 ccagcatggtgctcagcgccatgggtgtatgtatggagcctccaccaagccccagt 1044

RESULT 9
AAI29477 standard; cDNA; 518 BP.
XX AAI29477;
AC AAI29477;
DT 12-OCT-2001 (first entry)
XX Colon tumour related determined cDNA sequence for clone R0098.F12.
XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.
PN WO200149716-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000WO-US35596.
XX 30-DEC-1999; 93US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX (CORI-) CORIXA CORP.
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.
XX Colon tumor associated proteins and nucleic acids useful for the
XX prevention, diagnosis and treatment of colonic cancer -
XX
XX Claim 2; Page 415; 472pp; English.
XX
XX The present invention describes colon tumor associated proteins (I) and
XX the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX (II) may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate colon tumor associated protein (TCAP)
XX expression, such as colonic cancer. For example, (I) and (II) may be
XX used to treat disorders associated with decreased expression by
XX rectifying mutations or deletions in a patient's genome that affect the
XX activity of TCAPs by expressing inactive proteins or to supplement the
XX patients own production of them. Additionally, (II) may be used to
XX produce the TCAP proteins, by inserting the nucleic acids into a host
XX cell culturing the cell to express the protein. (II) and its
XX complementary sequences may also be used as DNA probes in diagnostic
XX polymerase chain reaction (PCR) and hybridisation assays to detect and
XX quantitate the presence of similar nucleic acids in samples, and
XX therefore which patients may be in need of restorative therapy. (I) may
XX also be used as antigens in the production of antibodies against TCAPs
XX and in assays to identify modulators of TCAP expression and activity.
XX Anti-(I) antibodies and antagonists may also be used to down regulate
XX TCAP expression and activity. The anti-(I) antibodies may also be used
XX as diagnostic agents for detecting the presence of TCAPs in samples
XX (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
XX and AA24494 to AA24523 represent nucleotide and amino acid sequences
XX given in the exemplification of the present invention.
XX
XX Sequence 518 BP; 123 A; 136 C; 149 G; 109 T; 1 other;

Query Match 30.7%; Score 365; DB 22; Length 518;
Best Local Similarity 99.4%; Pred. No. 4.5e-169;
Matches 515; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 200 CCTGGGTGGTGGAGCGATGGCGATTCACCGGATCCTGGAGCCCTGTTGAACATCC 259
DB 1 cctg99g9gtg9agcgaatg9gcgattccaccggaatcctg9gcttgaacatcc 60
QY 260 TCATCCCTGTGTTAGCCGATCGATATGTGCAGAGTCTCAAGGAATTCATCAACG 319
DB 61 tcatccctgtgtagacggaatcgatgtgcagagttctcaagaaattcatcaacg 120
QY 320 TGCTGAGCACTCGGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGGATGCTTT 379
DB 121 tgctgagcagtcggtgtgactctcgacaatgtaaactctgcaatcgatgagtccttt 180
QY 380 ACCTGGCGATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGATGCGC 439
DB 181 accg9cgcatactgagcccttaccaggaagcagtcaggtgtg9gagccctgagtcg9c 240
QY 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGCAAA 499
DB 241 tcacccagctagctcaaaacacatgagatcagagctcgcaactctcttgacaaag 300
QY 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATGTGAGATGCCATCAACCAAGCTGCTG 559
DB 301 tcttccgggaacggagtcctgtaatg9ccagcattgtgagtcgcatcaacccaagtcgctg 360
QY 560 ACTGCTGGGTATCCGCTGCTTCGCTTATGAGATCAGAGTATCCATGTCACCCCGG 619
DB 361 actgctg9g9tatccgctgcttcctc9tktatgagatcaaggaatccatgtg9ccacccg9g 420
QY 620 TGAAGAGTCTATCAGATGCGAGTGGAGGCGAGCGCGGAAACGGGCCACAGTTCCTAG 679
DB 421 tgaagagctctatgagatgcangtggaggcagagcggcggaacggccacagttctag 480
QY 680 ACTCTAGGGGACCGGAGAGTCGGCCATCAATGTGGCA 717

DB 481 agtctgagggagcccgagagtcggccatcaatgtggca 518
RESULT 10
AAK93925/C
XX ID AAK93925 standard; cDNA; 566 BP.
XX AAK93925;
XX AC
XX DT 06-NOV-2001 (first entry)
XX Human cDNA clone representative sequence, SEQ ID NO: 2385.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-0114089.
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX (HELI-) HELIX RES INST.
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX Example 11; SEQ ID NO 2385; 1380pp + sequence listing; English.
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
XX molecules have been determined. Primers for synthesising the full
XX length cDNA are useful for clarifying the function of the protein encoded by
XX the cDNA. The full length clones were obtained by construction of full
XX length enriched cDNA libraries that were synthesised by the oligo-capping
XX method. The primers enable the production of the full length cDNA easily
XX without any special methods. The present sequence was used as the
XX representative sequence from a human clone which was used in
XX homology searches to identify the clone.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in CD-ROM format directly from EPO.
XX Sequence 566 BP; 120 A; 152 C; 139 G; 155 T; 0 other;
Query Match 28.2%; Score 335; DB 22; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.4e-154;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 764 CTGAACAGATAAATCAGCAGCAGGAGCCAGTTCGCGCAAGCCCAAGGCTA 823
DB 517 CTGAACAGATAAATCAGCAGCAGGAGCCAGTTCGCGCAAGCCCAAGGCTA 458
QY 824 AAGCTGAAGCTATTCGATCTCGTTCGAGCTGTGACACACATATGAGATGACGAG 883
DB 457 AAGCTGAAGCTATTCGATCTCGTTCGAGCTGTGACACACATATGAGATGACGAG 398
QY 884 CTTCACTGACTGTGGCGGAGCAGTATGTCAGCGGTTCTCCAAACTGGCCAAAGACTCCA 943
DB 397 CTTCACTGACTGTGGCGGAGCAGTATGTCAGCGGTTCTCCAAACTGGCCAAAGACTCCA 338

```
QY 944 ACACATATCTACTGCCCTCCAAACCCCTGGCGATGTCACACAGCATGTGGCTCAGGCCATGG 1003
Db 337 ACACATATCTACTGCCCTCCAAACCCCTGGCGATGTCACACAGCATGTGGCTCAGGCCATGG 278
QY 1004 GTGTATATGGAGCCCTCACCRAAGCCCGAGTCCAGGAGCTCCAGACTCTCCAGTG 1063
Db 277 GTGTATATGGAGCCCTCACCRAAGCCCGAGTCCAGGAGCTCCAGACTCTCCAGTG 218
QY 1064 GGAGCAGCAGAGATGTCACGAGGTACAGATGCAAGT 1098
Db 217 GGAGCAGCAGAGATGTCACGAGGTACAGATGCAAGT 183

RESULT 11
ID AAS58752 standard; cDNA; 691 BP.
XX
AC AAS58752;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #1428 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-US9246.
XX
PR 24-MAR-2000; 2000US-191597P.
PR 04-MAY-2000; 2000US-202024P.
PR 05-MAY-2000; 2000US-202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
DR WPI; 2001-611627/70.
XX
PT New colon tumour proteins and related nucleic acid, useful for
PT treatment, prevention, diagnosis and monitoring of cancer
XX
PS Claim 4; Page 281; 299pp; English.
XX
CC Th present invention relates to the isolation of novel cDNA sequences
CC encoding for at least an immunogenic portion of human colon tumour
CC proteins. The sequences of the invention are useful in pharmaceutical
CC compositions and vaccines for the prevention and treatment of cancers
CC such as colon cancer. They are also useful for the diagnosis and
CC monitoring of such cancers. Antibodies to the colon tumour proteins
CC and antigen presenting cells that express polynucleotides encoding
CC colon tumour proteins can be used to inhibit the development of
CC cancers. T-cells that react specifically with colon tumour proteins
CC are useful for removing tumour cells from samples (e.g. blood) and
CC for cancer treatment. The polynucleotides sequences are also useful in
CC gene therapy. AAS57325-AAS58840 represent the cDNA sequences of the
CC invention that encode for portions of human colon tumour proteins.
XX
SQ Sequence 691 BP; 183 A; 168 C; 184 G; 140 T; 16 other;

Query Match 27.9%; Score 331; DB 23; Length 691;
Best Local Similarity 99.7%; Pred. No. 2.2e-152;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 200 CTGGGTGTCGAGCGAATGGCCGATTCACCGGATCCTGGAGCTGCTTCACATCC 259
Db 1 cctgggtggagcgaaatggccgatccacccgagctgagcctggagctggaacatcc 60
```

```
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATGTCATCAACG 319
Db 61 tcatccctgtgttagaccggatccgatatgtcagagctctcaaggaaattgcatcaacg 120
QY 320 TGCCTGAGCAGTCGCTGTGACTCTCGACAATGTAAGTCTGCAAAATCGATGGAGTCTTT 379
Db 121 tgcttgagcagtcgctgtgactctcgacaatgtaactctgcaaatcgatggagtccttt 180
QY 380 ACTCTGCGCATATGACCCCTTACAGGCAAGCTACCGTGTGGAGGACCTGAGTATGCGCG 439
Db 181 acctgcgatcatggaaccttaagaaggaagctacggtgtggaggaacctgagatgcgcg 240
QY 440 TCACCCAGCTAGTCTCAACACACATGAGATCAGAGTCGCGCAAACTCTCTNFGACAAAG 499
Db 241 tcaccagctagctcaaacacacatgagatcagagctcggaactctctctggacaaag 300
QY 500 TCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTCGATGCCATCAACCAAGCTGCTG 559
Db 301 tcttcgggaaaggagtcctcgaaatgccagcatttgatgccatcaaccaagctgctg 360
QY 560 ACTGCTGGGGTATCCGCTGCCT 581
Db 361 actgctggggtatccgctgcct 382

RESULT 12
AAK94818
ID AAK94818 standard; cDNA; 2064 BP.
XX
AC AAK94818;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3958.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation
XX
PS Claim 8; SEQ ID NO 3958; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
```



```
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 2064 BP; 479 A; 506 C; 581 G; 498 T; 0 other;

Query Match      24.0%; Score 285; DB 22; Length 2064;
Best Local Similarity 100.0%; Pred. No. 8.4e-130;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 583 CGTTATGATGATCAAGGATATCCATGTCGCCACCCGCGGTGAAGAGTCTATGCAGATGCAG 642
Dy 1169 cgttatgagatcaaggatatccatcgtgccaccgccggtgaaagagttctatgcagatgcag 1228
Qy 643 GTGGAGGACAGAGCGCGGAACGGCGCCACAGTTCTAGAGTCTGAGGGACCCGAGAGTCG 702
Dy 1229 gtgaggcagagcgcggaacggcgccacagttctagagttcgtgagggaccgcgagtcg 1288
Qy 703 GCCATCAATGTGGCAGAGGGAAGAACAGCCAGCCAGATCCTGGCCTCGGAAGCAGAAAG 762
Dy 1289 gccatcaatgtggcagaagggaagaaacaggccagatcctggcctccgaagcagaaag 1348
Qy 763 GCTGAACAGATAAATCAGGCAGCAGGAGGAGCCAGTCTGCGCAAGCCCAAGGCT 822
Dy 1349 gctgaacagataaatacaggcagcagggagggcagttctggtggcgaaggcccaaggct 1408
Qy 823 AAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACACAAACAT 867
Dy 1409 aaagctgaagctatttcgaatcctggctgctgcagctctgcacacaacat 1453

RESULT 13
AAL23495/c
ID AAL23495 standard; cDNA; 313 BP.
XX
AC AAL23495;
XX
DT 07-DEC-2001 (first entry)
DE Human breast cancer expressed polynucleotide 15952.
KW Human; breast cancer; cell marker; cytostatic; ss.
XX Homo sapiens.
OS
PN WO200151628-A2.
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
XX Claim 1; Page 2911; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
```

```
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 313 BP; 60 A; 86 C; 89 G; 78 T; 0 other;

Query Match      21.3%; Score 253; DB 22; Length 313;
Best Local Similarity 99.7%; Pred. No. 4.6e-114;
Matches 303; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 776 ATCAGGAGCAGGAGAGAGCCAGTTCAGTCTTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTA 835
Dy 313 ATCAGGAGCAGGAGAGAGCCAGTTCAGTCTTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTA 254
Qy 836 TTGCAATCCTGGCTGCAGCTCTGACACAACATAATGGAGATGCAGCAGCTTTCACCTGACTG 895
Dy 253 TTGCAATCCTGGTTGCAGCTCTGACACAACATAATGGAGATGCAGCAGCTTTCACCTGACTG 194
Qy 896 TGGCCGAGCAGTATGTACAGCGCTTCTCCAACTGGCCCAAGGACTCCAACTATCTCTAC 955
Dy 193 TGGCCGAGCAGTATGTACAGCGCTTCTCCAACTGGCCCAAGGACTCCAACTATCTCTAC 134
Qy 956 TGCCCTCCAAACCCCTGGCGATGTCACACAGATGTTGGCTCAGGCCATGGTGTATATGGAG 1015
Dy 133 TGCCCTCCAAACCCCTGGCGATGTCACACAGATGTTGGCTCAGGCCATGGTGTATATGGAG 74
Qy 1016 CCCTCACCAAGGCCCGAGTGCAGGAGTCCAGACTCCTCCAGTGGGAGCAGCAGAG 1075
Dy 73 CCCTCACCAAGGCCCGAGTGCAGGAGTCCAGACTCCTCCAGTGGGAGCAGCAGAG 14
Qy 1076 ATGT 1079
Dy 13 ATGT 10

RESULT 14
AAS46972/c
ID AAS46972 standard; cDNA; 301 BP.
XX
AC AAS46972;
XX
DT 18-DEC-2001 (first entry)
DE Human breast cancer cDNA clone JBTT21.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy.
XX Homo sapiens.
OS
PN WO200179286-A2.
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US12164.
XX
PR 17-APR-2000; 2000US-0551621.
PR 08-JUN-2000; 2000US-0590751.
PR 22-JUN-2000; 2000US-0604287.
PR 20-JUL-2000; 2000US-0620405.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX
XX Breast Tumour Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer -
XX Claim 1; Page 155; 297pp; English.
PS
```


XX The invention relates to isolated breast tumour proteins and
CC nucleic acids that encode them, including immunogenic fragments of the
CC proteins. Also included are expression vectors expressing the
CC proteins, transformed cells and antibodies raised against the proteins or
CC an antigen presenting cell expressing the protein. The proteins and
CC nucleic acids may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate breast tumour protein expression,
CC i.e. breast tumours and breast cancer e.g. by gene therapy. The nucleic
CC acids and their complements may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The proteins, nucleic acids and antibodies may be used in assays
CC to identify modulators (e.g. antagonists) of breast tumour protein
CC expression and activity. The antibodies and antagonists may also be used
CC to down regulate expression and activity. The antibodies may also be used
CC as diagnostic agents for detecting the presence of the proteins in
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other
CC immuno-purification diagnostic techniques. The present sequence is
CC a cDNA from a breast tumour cDNA library isolated by subtractive
CC hybridisation against a normal breast cDNA library.
XX
SQ Sequence 301 BP; 48 A; 96 C; 72 G; 85 T; 0 other;

Query Match 21.0%; Score 250; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 1.4e-112;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 626 AGTCTATGCAGATCGAGTGGAGCGCGGAAACGCGCCACAGTCTTAGAGTCTG 685
DB 301 AGTCTATGCAGATCGAGTGGAGCGCGGAAACGCGCCACAGTCTTAGAGTCTG 242
QY 686 AGGGACCCGAGAGTCGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCCAGATCCTGG 745
DB 241 AGGGACCCGAGAGTCGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCCAGATCCTGG 182
QY 746 CCTCCGAAGCAAGAGCTGAACAGATAAATCAGGCAGCAGGAGGCCAGTGCAGTTC 805
DB 181 CCTCCGAAGCAAGAGCTGAACAGATAAATCAGGCAGCAGGAGGCCAGTGCAGTTC 122
QY 806 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 865
DB 121 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 62
QY 866 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 925
DB 61 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 2
QY 926 A 926
DB 1 A 1

RESULT 15
AAFL17542/c
ID AAF17542 standard; cDNA; 301 BP.
XX
AC AAF17542;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human breast cancer associated JBT21 coding sequence.
XX
KW Human; breast cancer associated gene; vaccine; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200060076-A2.
XX
PD 12-OCT-2000.
XX
PF 15-FEB-2000; 2000WO-US05308.

XX 02-APR-1999; 99US-0285480.
PR 23-JUN-1999; 99US-0339338.
PR 02-SEP-1999; 99US-0389681.
PR 03-NOV-1999; 99US-0433826.
XX (CORI-) CORIXA CORP.
PA Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2001-122627/13.
DR
XX An isolated polypeptide useful for the treatment and diagnosis of
PT tumors e.g. breast cancer comprises at least an immunogenic portion of
PT a breast tumor protein.
XX
PS Claim 6; Page 95; 238pp; English.
XX
CC The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumours.
XX
SQ Sequence 301 BP; 48 A; 96 C; 72 G; 85 T; 0 other;

Query Match 21.0%; Score 250; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 1.4e-112;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 626 AGTCTATGCAGATCGAGTGGAGCGCGGAAACGCGCCACAGTCTTAGAGTCTG 685
DB 301 AGTCTATGCAGATCGAGTGGAGCGCGGAAACGCGCCACAGTCTTAGAGTCTG 242
QY 686 AGGGACCCGAGAGTCGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCCAGATCCTGG 745
DB 241 AGGGACCCGAGAGTCGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCCAGATCCTGG 182
QY 746 CCTCCGAAGCAAGAGCTGAACAGATAAATCAGGCAGCAGGAGGCCAGTGCAGTTC 805
DB 181 CCTCCGAAGCAAGAGCTGAACAGATAAATCAGGCAGCAGGAGGCCAGTGCAGTTC 122
QY 806 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 865
DB 121 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 62
QY 866 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 925
DB 61 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 2
QY 926 A 926
DB 1 A 1

Search completed: September 22, 2002, 17:25:55
Job time: 4267 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:26:33 ; Search time 50.25 Seconds
(without alignments)
5807.216 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GGCTTCTGGAGCACCCT.....GGAAGCAGATTTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	1	US-08-781-562-2
2	21	1.8	28958	1	US-08-258-261B-6
3	21	1.8	28958	1	US-08-456-837-6
4	21	1.8	28958	1	US-08-457-342-6
5	21	1.8	28958	1	US-08-457-646A-6
6	21	1.8	28958	1	US-08-458-076A-6
7	21	1.8	28958	1	US-08-764-233A-4
8	21	1.8	28958	1	US-08-457-335A-6
9	21	1.8	28958	1	US-08-729-214-6
10	21	1.8	28958	3	US-09-028-934-6
11	21	1.8	49377	1	US-08-764-233A-1
12	18	1.5	2371	2	US-08-343-443B-1
13	18	1.5	2412	1	US-08-437-027-18
14	17	1.4	576	1	US-08-086-428B-34
15	17	1.4	576	2	US-08-468-570-34
16	17	1.4	576	2	US-08-290-665A-34
17	17	1.4	576	5	PCT-US95-10398-34
18	17	1.4	603	4	US-09-328-111-244
19	17	1.4	844	4	US-08-998-416-349
20	17	1.4	1025	2	US-08-482-728A-20
21	17	1.4	1468	2	US-09-074-512-2
22	17	1.4	1721	3	US-09-173-581-16
23	17	1.4	1721	4	US-09-420-915-16
24	17	1.4	1785	2	US-08-559-505-3
25	17	1.4	1785	2	US-08-749-907-3
26	17	1.4	1785	4	US-09-241-581B-7
27	17	1.4	1785	5	PCT-US95-07721-7

28	17	1.4	2060	1	US-08-480-547A-11	Sequence 11, Appl
29	17	1.4	2060	1	US-08-250-847B-11	Sequence 11, Appl
30	17	1.4	2060	2	US-08-463-949A-11	Sequence 11, Appl
31	17	1.4	2060	3	US-08-464-410A-11	Sequence 11, Appl
32	17	1.4	2060	5	PCT-US94-06066-11	Sequence 11, Appl
c 33	17	1.4	2311	4	US-08-976-239-26	Sequence 26, Appl
c 34	17	1.4	35060	3	US-08-814-095-7	Sequence 7, Appl
35	16	1.3	266	1	US-08-594-031-32	Sequence 32, Appl
36	16	1.3	266	1	US-08-594-031-124	Sequence 23, Appl
37	16	1.3	574	4	US-09-073-297-23	Sequence 124, App
c 38	16	1.3	720	4	US-08-998-416-992	Sequence 992, App
39	16	1.3	876	4	US-08-818-112-19	Sequence 19, Appl
40	16	1.3	876	4	US-08-818-111-19	Sequence 19, Appl
41	16	1.3	876	4	US-09-056-556-19	Sequence 19, Appl
42	16	1.3	977	1	US-08-017-522A-1	Sequence 1, Appl
43	16	1.3	977	6	5215895-2	Patent No. 5215895
44	16	1.3	1018	1	US-08-444-083-7	Sequence 7, Appl
45	16	1.3	1018	1	US-08-286-304-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-781-562-2
; Sequence 2, Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.562
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/POCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1188 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-08-781-562-2

Query Match 100.0%; Score 1188; DB 1; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

	Matches	1188:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
Qy	1	GGCTTCTGGGAGCNACCGCTCCGCTCGTCTCGTTGGTTCCGGAGGTCGCTCGCGGGTGG	60							
Db	1	GGCTTCTGGGAGCNACCGCTCCGCTCGTCTCGTTGGTTCCGGAGGTCGCTCGCGGGTGG	60							
Qy	61	GAATGCTGGCGGGCGGGCGGGGGGACATGGGGGCCCTTTTGTGTGAGGGGCTCTCTACT	120							
Db	61	GAATGCTGGCGGGCGGGCGGGGGGACATGGGGGCCCTTTTGTGTGAGGGGCTCTCTACT	120							
Qy	121	GGCTTCTGGCGGGCGGGCGGGGCCCTCTCTTGATTTGCCCCGAAACACCGTGGTACTG	180							
Db	121	GGCTTCTGGCGGGCGGGCGGGGCCCTCTCTTGATTTGCCCCGAAACACCGTGGTACTG	180							
Qy	181	TTCTGTCGCGACGAGAGGCCCTGGGTGGTGAGCGCAATGGGCCGATTCACCGGATCTCTG	240							
Db	181	TTCTGTCGCGACGAGAGGCCCTGGGTGGTGAGCGCAATGGGCCGATTCACCGGATCTCTG	240							
Qy	241	GAGCCTGGTTGAACATCCTCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTC	300							
Db	241	GAGCCTGGTTGAACATCCTCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTC	300							
Qy	301	AAGGAAATGTTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAAATGTAACCTG	360							
Db	301	AAGGAAATGTTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAAATGTAACCTG	360							
Qy	361	CAAAATCGATGGAGTCCTTTTACTCTGGCATCATGACCCCTTACAAGCGAAGCTACGGTCTG	420							
Db	361	CAAAATCGATGGAGTCCTTTTACTCTGGCATCATGACCCCTTACAAGCGAAGCTACGGTCTG	420							
Qy	421	GAGGACCTGTAGTATGCGCGTCAACCAGTGTAGCTCAAAACCAACCATGAGATCGGCTCGC	480							
Db	421	GAGGACCTGTAGTATGCGCGTCAACCAGTGTAGCTCAAAACCAACCATGAGATCGGCTCGC	480							
Qy	481	AAATCTCTNTGGACAAGTCTTCCGGGAAACGGAGTCCCTGAATGCCAGCATTTGTGGAT	540							
Db	481	AAATCTCTNTGGACAAGTCTTCCGGGAAACGGAGTCCCTGAATGCCAGCATTTGTGGAT	540							
Qy	541	GCCATCAACCAAGCTGTGACTGTCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGAT	600							
Db	541	GCCATCAACCAAGCTGTGACTGTCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGAT	600							
Qy	601	ATCCATGTGCCACCCCGGGTGAAGAGTCTATGCAGATGCAGGTGGAGCGAGCGCGGG	660							
Db	601	ATCCATGTGCCACCCCGGGTGAAGAGTCTATGCAGATGCAGGTGGAGCGAGCGCGGG	660							
Qy	661	AAACGGGCCACAGTCTAGAGTCTGAGGGGACCGAGAGTCGGCCATCAATGTGGCAGAA	720							
Db	661	AAACGGGCCACAGTCTAGAGTCTGAGGGGACCGAGAGTCGGCCATCAATGTGGCAGAA	720							
Qy	721	GGGAAGAAACAGGCCAGATCTCGGCCCTCCGAACGAGAAAGGCTGAACAGATAAATCAG	780							
Db	721	GGGAAGAAACAGGCCAGATCTCGGCCCTCCGAACGAGAAAGGCTGAACAGATAAATCAG	780							
Qy	781	GCACGAGGAGGCCAGTGTGAGTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTCTGA	840							
Db	781	GCACGAGGAGGCCAGTGTGAGTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTCTGA	840							
Qy	841	ATCTTGGCTGCAGCTCTGACACACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCC	900							
Db	841	ATCTTGGCTGCAGCTCTGACACACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCC	900							
Qy	901	GAGCAGTATGTCAGCGGGTCTCCAAACTGGCCAAAGGACTCCAACACTATCTACTGCC	960							
Db	901	GAGCAGTATGTCAGCGGGTCTCCAAACTGGCCAAAGGACTCCAACACTATCTACTGCC	960							
Qy	961	TCCAACCCCTGGCGATGTCAACAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTC	1020							
Db	961	TCCAACCCCTGGCGATGTCAACAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTC	1020							
Qy	1021	ACCAAGCCCGAGTGCAGGGACTTCCAGACTCTCTCTCCAGTGGGAGCAGAGATGTC	1080							
Db	1021	ACCAAGCCCGAGTGCAGGGACTTCCAGACTCTCTCTCCAGTGGGAGCAGAGATGTC	1080							

```

Qy 1081 CAGGGTACAGATCAAGTNTGATGAGGAACCTTGATCGAGTCAAGATGAGTTCAGTGGAGC 1140
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CAGGGTACAGATCAAGTNTGATGAGGAACCTTGATCGAGTCAAGATGAGTTCAGTGGAGC 1140

Qy 1141 TGGGCTTNGCCAGGAGTCTGGGGACAAGGAAGCAGATTTTCCTCATTT 1188
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TGGGCTTNGCCAGGAGTCTGGGGACAAGGAAGCAGATTTTCCTCATTT 1188

RESULT 2
US-08-258-261B-6
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaifney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NV
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-258-261B-6

```

```
RESULT 3
US-08-456-837-6
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456.837
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-456-837-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCCGCCGCGCTC 149
Db 23859 GCCGCGCTCCGCCGCGCTC 23879
|||||
RESULT 4
US-08-457-342-6
; Sequence 6, Application US/08457342
; Patent No. 5662898
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
```

```
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,342
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-457-342-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCCGCCGCGCTC 149
Db 23859 GCCGCGCTCCGCCGCGCTC 23879
|||||
RESULT 5
US-08-457-646A-6
; Sequence 6, Application US/08457646A
; Patent No. 5679560
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
```

APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 GCGCGCTCCGGCGCGCTC 149
Db 23859 GCGCGCTCCGGCGCGCTC 23879

RESULT 6
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive

CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-458-076A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 GCGCGCTCCGGCGCGCTC 149
Db 23859 GCGCGCTCCGGCGCGCTC 23879

RESULT 7
US-08-764-233A-4
Sequence 4, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/764,233A

;
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-335A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCGCGCTCCGCGCGCCTC 149
|||||
DB 23859 GCGCGCTCCGCGCGCCTC 23879

RESULT 8
US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent No. 5723759
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,335A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205

;
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-335A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCGCGCTCCGCGCGCCTC 149
|||||
DB 23859 GCGCGCTCCGCGCGCCTC 23879

RESULT 9
US-08-729-214-6
; Sequence 6, Application US/08729214
; Patent No. 5817502
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,214
; FILING DATE: TBA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-729-214-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCGGCGCGCCTC 149
|||||
DB 23859 GCCGCGCTCGGCGCGCCTC 23879

RESULT 10
US-09-028-934-6
; Sequence 6, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrolyttrin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028.934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 36,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-028-934-6

Query Match 1.8%; Score 21; DB 3; Length 28958;

Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 GCCGCGCTCGGCGCGCCTC 149
|||||
DB 23859 GCCGCGCTCGGCGCGCCTC 23879

RESULT 11
US-08-764-233A-1
; Sequence 1, Application US/08764233A
; Patent No. 5716849
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764.233A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 36,241
; REFERENCE/DOCKET NUMBER: 1506/CIP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Sorangium cellulosum
; IMMEDIATE SOURCE:
; CLONE: p98/1, pJL3, and pVRM15
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 383..760
; OTHER INFORMATION: /product= "SorR"
; OTHER INFORMATION: /note= "This gene encodes a protein that is highly homolog
; OTHER INFORMATION: the reductase domains of type I PKSs such as eryA from
; OTHER INFORMATION: Saccharopolyspora erythraea."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 927..19874
; OTHER INFORMATION: /product= "SorA"
; OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKSs t

OTHER INFORMATION: are known to be involved in the synthesis of polyketide
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS genes"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "Sorm"
OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly homologous to the methyltransferase from Streptomyces hygroscopicus that is involved in the synthesis of the polyketide rapamycin."
US-08-764-233A-1
Query Match 1.8%; Score 21; DB 1; Length 49377;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 GCCGGCTCCGGCGGCCTC 149
Db 40162 GCCGGCTCCGGCGGCCTC 40182
RESULT 12
US-08-343-443B-1/c
Sequence 1, Application US/0834343B
Patent No. 5968734
GENERAL INFORMATION:
APPLICANT: Aurias, Alain
APPLICANT: Delattre, Olivier
APPLICANT: Desmaze, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Peter, Martine
APPLICANT: Ploougaestel, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF

TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2371 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1992
US-08-343-443B-1
Query Match 1.5%; Score 18; DB 2; Length 2371;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 194 AGGAGCCCTGGGTGGTG 211
Db 358 AGGAGCCCTGGGTGGTG 341
RESULT 13
US-08-437-027-18/c
Sequence 18, Application US/08437027
Patent No. 5670317
GENERAL INFORMATION:
APPLICANT: Landanyi, Marc
APPLICANT: Gerald, William
TITLE OF INVENTION: A DIAGNOSTIC TEST FOR TEST FOR THE DESMOPLASTIC
TITLE OF INVENTION: SMALL ROUND CELL TUMOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: 194
FILING DATE: US/08/437-027
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 46416/JPM/CCA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2412 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-437-027-18

Query Match 1.5%; Score 18; DB 1; Length 2412;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AGGAGCGCTGGTGGTGG 211
DB 358 AGGAGCGCTGGTGGTGG 341

RESULT 14
US-08-086-428B-34
Sequence 34, Application US/08086428B
Patent No. 5514539
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086.428B
FILING DATE: 29-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 2412 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S83
US-08-437-027-18

Query Match 1.5%; Score 18; DB 1; Length 2412;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S83
US-08-086-428B-34

Query Match 1.4%; Score 17; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CTGCTGACTGCTGGGGT 570
DB 553 CTGCTGACTGCTGGGGT 569

RESULT 15
US-08-468-570-34
Sequence 34, Application US/08468570
Patent No. 5871962
GENERAL INFORMATION:
APPLICANT: BUKH, J., MILLER, R.H. AND
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468.570
FILING DATE: 6-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/086.428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S83
US-08-468-570-34

Query Match 1.4%; Score 17; DB 2; Length 576;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches	639:	Conservative	0:	Mismatches	2:	Indels	0:	Gaps	0:
QY 264	CCCTGTGTAGACCGGATCCGATATGTGACAGAGTCTCAAGAAATTTGCATCAACGTGCC								323
Db 2	CCCTGTGTAGACCGGATCCGATATGTGACAGAGTCTCAAGAAATTTGCATCAACGTGCC								61
QY 324	TGACAGTCCGCTGTGACTCTCGACATGTAACTCTGCAATTCGAGGAGTCTTTACCT								383
Db 62	TGACAGTCCGCTGTGACTCTCGACATGTAACTCTGCAATTCGAGGAGTCTTTACCT								121
QY 384	CGGATCATGAGCCCTTCAAGGAGGAGTACGGTGTGAGAGCCCTGAGTATGGCGTAC								443
Db 122	CGGATCATGAGCCCTTCAAGGAGGAGTACGGTGTGAGAGCCCTGAGTATGGCGTAC								181
QY 444	CCAGCTAGCTCAAAACCAACCATGAGATCAGAGCTGGCAAACTCTGTTGAGCAAAAGTCT								503
Db 182	CCAGCTAGCTCAAAACCAACCATGAGATCAGAGCTGGCAAACTCTGTTGAGCAAAAGTCT								241
QY 504	CCGGAAACGGAGTCCCTGGAATGCCAGATTGTGATGCCATCAACAGAGTCTGACTG								563
Db 242	CCGGAAACGGAGTCCCTGGAATGCCAGATTGTGATGCCATCAACAGAGTCTGACTG								301
QY 564	CTGGGATATCCGCTGCGCTTCGCTTATGAGATCAGAGATTCATGTCGCCACCCGGGTGA								623
Db 302	CTGGGATATCCGCTGCGCTTCGCTTATGAGATCAGAGATTCATGTCGCCACCCGGGTGA								361
QY 624	AGACTATATGAGATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG								683
Db 362	AGACTATATGAGATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG								421
QY 684	TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG								743
Db 422	TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG								481
QY 744	GGCTCCGAAAGCAGAAAAGGCTGAACAGATTAATCAGGAGGAGGAGGAGGAGGAGGAGG								803
Db 482	GGCTCCGAAAGCAGAAAAGGCTGAACAGATTAATCAGGAGGAGGAGGAGGAGGAGGAGG								541
QY 804	TCTGGCAGAAAGGCTGAAGCTGAAGCTATTGCAATCTGAGCTGAGGCTCTGACACA								863
Db 542	TCTGGCAGAAAGGCTGAAGCTGAAGCTATTGCAATCTGAGCTGAGGCTCTGACACA								601
QY 864	ACATATGAGAGATCAGACAGCTTCACTGACTGTGGCCGAGC								904
Db 602	ACATATGAGAGATCAGACAGCTTCACTGACTGTGGCCGAGC								642
RESULT 15									
LOCUS	BM471641	1037 bp	mRNA	Linear	EST 05-FEB-2002				
DEFINITION	AGENCOURT_6465267 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539088								
ACCESSION	BM471641								
VERSION	BM471641.1	GI:18520683							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:								
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.								
TITLE	NIH-MGC http://mgc.nci.nih.gov/								
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)								
COMMENT	Unpublished (1999)								
	Contact: Robert Strausberg, Ph. D.								
	Email: cgaabs-remail.nih.gov								
	Tissue Procurement: ATCC/DCID/DBP								
	cDNA Library Preparation: Life Technologies Inc.								
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)								
	DNA Sequencing by: Agencourt Bioscience Corporation								
	Clone distribution: MGC clone distribution information can be								
	found through the I.M.A.G.E. Consortium/LNL at:								
	http://image.llnl.gov								
	Plate: L1AM12233	row: a	column: 09						

FEATURES	High quality sequence stop: 589.
source	Location/Qualifiers
	1..1037
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5539088"
	/clone_id="NIH_MGC_72"
	/tissue_type="melanotic melanoma"
	/lab_host="DH10B (phage-resistant)"
	/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT	236 a 321 c 251 g 229 t
ORIGIN	
Query Match	45.1%; Score 536; DB 10; Length 1037;
Best Local Similarity	99.7%; Pred. No. 3.7e-259;
Matches	636; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 175	GTACTGTTGCGCCGACAGAGGCTGGTGTGAGAGCCGAATGGCCATTCACCGG
Db 1	GTACTGTTGCGCCGACAGAGGCTGGTGTGAGAGCCGAATGGCCATTCACCGG
QY 235	ATTCGAGACCTGTTGTAACATCCTCATCCCTGTTAGACCGGATCGATTATGTGAG
Db 61	ATTCGAGACCTGTTGTAACATCCTCATCCCTGTTAGACCGGATCGATTATGTGAG
QY 295	AGTCTCAAGAAATTTGTCATCAACGCTGCTGAGCAGATGCGCTGTGACTCTGACAAATGA
Db 121	AGTCTCAAGAAATTTGTCATCAACGCTGCTGAGCAGATGCGCTGTGACTCTGACAAATGA
QY 355	ACTCTGCAAAATGATGAGTGCCTTACTCGGCAATCATGACACCTTAAAGGCAAGCTAC
Db 181	ACTCTGCAAAATGATGAGTGCCTTACTCGGCAATCATGACACCTTAAAGGCAAGCTAC
QY 415	GGTGTGAGGAGACCTGAGATATGACCGTACCCAGTATGCTCAACAAACCATGAGATCAGAG
Db 241	GGTGTGAGGAGACCTGAGATATGACCGTACCCAGTATGCTCAACAAACCATGAGATCAGAG
QY 475	CTCGGCAAACTCTCTGTCGAAAGTCTTCGGGAAACGGAGTCCCTGAATGCCAGATT
Db 301	CTCGGCAAACTCTCTGTCGAAAGTCTTCGGGAAACGGAGTCCCTGAATGCCAGATT
QY 535	GTGATGCCATCAACCAAGCTGCTGACTGCTGGGATATCCCTGCTTCGTTATGAGATC
Db 361	GTGATGCCATCAACCAAGCTGCTGACTGCTGGGATATCCCTGCTTCGTTATGAGATC
QY 595	AAGATATCCATGTGCGCACCCCGGCTGAAGAAGTCTATGACAGATGAGGTGAGGAGAG
Db 421	AAGATATCCATGTGCGCACCCCGGCTGAAGAAGTCTATGACAGATGAGGTGAGGAGAG
QY 655	CGGCGAAACGGGCGACAGTTCTAGAGCTGAGGAGGAGCCGAGAGTGGCCATCAATGTG
Db 481	CGGCGAAACGGGCGACAGTTCTAGAGCTGAGGAGGAGCCGAGAGTGGCCATCAATGTG
QY 715	GCAAGAAGGAGAAACAGGCCCATATCTGCGCTCCGAGAGGAGGAGGAGGAGGAGGAGGAG
Db 541	GCAAGAAGGAGAAACAGGCCCATATCTGCGCTCCGAGAGGAGGAGGAGGAGGAGGAGGAG
QY 775	AATCAGCAGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db 601	AATCAGCAGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

Search completed: September 22, 2002, 16:48:50
Job time: 5157 sec

```
Db 601 GCAGTATGTACGGCGTCTCCAAACTGGCCAGAGCATCTCCAACTATCTACTGCC 657
|||||
RESULT 13
Bg036463 749 bp mRNA linear EST 24-JAN-2001
LOCUS 60232409F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4427787 5',
DEFINITION mRNA sequence.
Bg036463
ACCESSION Bg036463.1 GI:12431652
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10177 row: a column: 04
High quality sequence stop: 708.
Location/Qualifiers
1. 749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4427787"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; site_1: NotI;
site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 190 a 183 c 222 g 154 t
ORIGIN
Query Match 46.5%; Score 552; DB 10; Length 749;
Best Local Similarity 99.8%; Pred. No. 2.9e-267;
Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 301 AGTGAGTTCTGGCCGAAGGCCAAGGCTAAGCTATTGATCTGGCTGACCT 360
|||||
Qy 856 CTGACACAACATTAATGAGATGCAGACCTTCACGTAGTGCGCCGACAGATGTACAG 915
|||||
Db 361 CTGACACAACATTAATGAGATGCAGACCTTCACGTAGTGCGCCGACAGATGTACAG 420
|||||
Qy 916 GCGTTCTCCAAACTGGCCAGAGACTCCAAACTATCTACTGCTCCCAACCTGGCGAT 975
|||||
Db 421 GCGTTCTCCAAACTGGCCAGAGACTCCAAACTATCTACTGCTCCCAACCTGGCGAT 480
|||||
Qy 976 GTGACACAGATGTGGTCTGAGCCCATGGGTGTATATGAGAGCCCTCCAAAGCCCACTG 1035
|||||
Db 481 GTGACACAGATGTGGTCTGAGCCCATGGGTGTATATGAGAGCCCTCCAAAGCCCACTG 540
|||||
Qy 1036 CCAGGACTCCAGACTCTCTCCAGTGGAGACAGACAGATGTCCAGAGGTACAGATGA 1095
|||||
Db 541 CCAGGACTCCAGACTCTCTCCAGTGGAGACAGACAGATGTCCAGAGGTACAGATGA 600
|||||
Qy 1096 AGT 1098
|||
Db 601 AGT 603
```

```
RESULT 14
Bg828070 770 bp mRNA linear EST 22-MAY-2001
LOCUS Bg828070
DEFINITION 60275343F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4905945 5',
mRNA sequence.
ACCESSION Bg828070
VERSION Bg828070.1 GI:14175657
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1808 row: d column: 10
High quality sequence stop: 767.
Location/Qualifiers
1. 770
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4905945"
/clone_lib="NIH_MGC_17"
/tissue_type="thabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; site_1: EcoRI;
site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(C). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT 200 a 206 c 216 g 148 t
ORIGIN
Query Match 45.4%; Score 539; DB 10; Length 770;
Best Local Similarity 99.7%; Pred. No. 1.1e-260;
```

	(Stratagene) and Superscript II RT (Life Technologies). "			
BASE COUNT	174 a	182 c	188 g	116 t
ORIGIN				

Query Match	47.08;	Score	558;	DB	10;	Length	660;
Best Local Similarity	99.78;	Pred.	No. 2.6e	270;			
Matches	658;	Conservative	0;	Mismatches	2;	Indels	0;
						Gaps	0;

Best Local Similarity 99.7%; Pred. NO. 2.6e-270;

Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY		394	GACCTTCAAGGCGCAGGTACGGTGTTGGAGGACCCTGAGTATGCCGTACCCAGGTACT	453
Db		1	GACCTTCAAGGCGCAGGTACGGTGTTGGAGGACCCTGAGTATGCCGTACCCAGGTACT	60
OY		454	CAAACCAACCATGAGATCAGAGGCTCGGCACAACTCTCTNTGGACAAGTCCTTCGGGAACGG	513
Db		61	CAAACCAACCATGAGATCAGAGGCTCGGCACAACTCTCTTGGACAAGTCCTTCGGGAACGG	120
OY		514	GAGTCCCTGAATTGCCAGCATTTGTGGATGCCATCAACCAAGCTGTGACTGCTGGGGTATC	573
Db		121	GAGTCCCTGAATTGCCAGCATTTGTGGATGCCATCAACCAAGCTGTGACTGCTGGGGTATC	180
OY		574	CGGTGCCCTNCGTTTATGAGATCAAGGATATATGCAATGTCACCCCGGGGTGAAGAAGCTATG	633
Db		181	CGGTGCCCTNCGTTTATGAGATCAAGGATATATGCAATGTCACCCCGGGGTGAAGAAGCTATG	240
OY		634	CAGATGCAAGTGTGAGAGGCGAGAGCGGGCGAAGCGGCACAGTTCTAGAGTCTGAGGGAGCC	693
Db		241	CAGATGCAAGTGTGAGAGGCGAGAGCGGGCGAAGCGGCACAGTTCTAGAGTCTGAGGGAGCC	300
OY		694	CGAGAGTCGCCCATTCATATGTGGCAGAGAGGGAAGAAACAGCCCCAGATCCTGGCCCTCCGA	753
Db		301	CGAGAGTCGCCCATTCATATGTGGCAGAGAGGGAAGAAACAGCCCCAGATCCTGGCCCTCCGA	360
OY		754	GCAGAAAAAGGCTGMAACACATATAATTCAGGCACAGAGAGGCGCAGTTCGTGGCGAAG	813
Db		361	GCAGAAAAAGGCTGMAACACATATAATTCAGGCACAGAGAGGCGCAGTTCGTGGCGAAG	420
OY		814	GCCAAGGCTTAAGAGCTGAGAGCTATTTCGATTCGTGGCTGAGCTCTACACACATATAATGA	873
Db		421	GCCAAGGCTTAAGAGCTGAGAGCTATTTCGATTCGTGGCTGAGCTCTACACACATATAATGA	480
OY		874	GATSCAGCAGGCTTACGACTGCTGGCCGAGCAGTAGATATGTCAGGCGCTTCTCCAACTGGCC	933
Db		481	GATSCAGCAGGCTTACGACTGCTGGCCGAGCAGTAGATATGTCAGGCGCTTCTCCAACTGGCC	540
OY		934	AAGGACTTCAACACTATCTTACTGCGCTTCCAAACCTGGCGATGTACCAAGCATGGTGGCT	993
Db		541	AAGGACTTCAACACTATCTTACTGCGCTTCCAAACCTGGCGATGTACCAAGCATGGTGGCT	600
OY		994	CAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAGCCCAGATGCCAGGGACATCCCAACTCA	1053
Db		601	CAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAGCCCAGATGCCAGGGACATCCCAACTCA	660
RESULT 12				
		Bt838402		
LOCUS		Bt838402	702 bp mRNA linear EST 04-OCT-2001	
DEFINITION		6030833k6F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5222655 5'		
ACCESSION		Bt838402	mRNA sequence.	
VERSION		Bt838402.1	GI:15949952	
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		NIH-MGC http://mgc.nci.nih.gov/		
AUTHORS		National Institutes of Health, Mammalian Gene Collection (MGC)		
TITLE		Unpublished (1999)		
JOURNAL		Contact: Robert Strausberg, Ph.D.		
COMMENT		Email: cgapds@emall.nih.gov		
		Tissue Procurement: Life Technologies, Inc.		

	CDNA Library Preparation: Life Technologies, Inc.
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LNL at:
	http://image.lnl.gov
	plate: LLM1559 row: p column: 16
	High quality sequence stop: 700.
FEATURES	
source	Location/Qualifiers
	1..702
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5222655"
	/clone_id="NIH_MGC_120"
	/lab_host="DH10B"
	/note="Organ: pooled pancreas and spleen; Vector:
	pCMV-Sport6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
	source anonymous pool of spleen and pancreas from 28 yo
	male. Library is oligo-dT primed and directionally cloned
	(EcoRV site is destroyed upon cloning). Average insert
	size 1.5 kb, insert size range 1-2.5 kb. Library is
	normalized and enriched for full-length clones and was
	constructed by C. Gruber (Invitrogen). Research Genetics
	tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT	183 a 187 c 199 g 133 t
ORIGIN	
Query Match	46.7%; Score 55; DB 10; Length 702;
Best Local Similarity	99.7%; Pred. No. 8,7e-269;
Matches 655; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	303 GGAAATTGTCATCAACGCTGAGCAGCTGGCTGTGACTCTCGACAATGTACTTGCA 362
Db	1 GGAATTTGTCAATCAACGGCTGAGCAGCTGGCTGTGACTCTCGACAATGTACTTGCA 60
OY	363 AATGATGAGATCCCTTAACCTGCCATATATGACCCTTACAAGGCACGATCGGTGGA 422
Db	61 AATCATGAGATCCCTTAACCTGCCATATATGACCCTTACAAGGCACGATCGGTGGA 120
OY	423 GGACCCCTGAGTAGCGCGCACCCAGCTAGCTCAACACCATGAGATGAGACTCGGCAA 482
Db	121 GGACCCCTGAGTAGCGCGCACCCAGCTAGCTCAACACCATGAGATGAGACTCGGCAA 180
OY	483 ACTCTCTNTGACAAAGTCTTCCGGGAAACGGAGTCCCTGTAATCCAGCATTTGTGATGC 542
Db	181 ACTCTCTGTGACAAAGTCTTCCGGGAAACGGAGTCCCTGTAATCCAGCATTTGTGATGC 240
OY	543 CATCAACCAAGCTGCTGACTCTGGGGTATCCCTGCCCTNCCTTATGATGATCAAGATAT 602
Db	241 CATCAACCAAGCTGCTGACTCTGGGGTATCCCTGCCCTNCCTTATGATGATCAAGATAT 300
OY	603 CCATGTGCCACCCGGGGAAGAAGTCTATGCAGTGAAGGTGGAGGCAGAGCGGGGAA 662
Db	301 CCATGTGCCACCCGGGGAAGAAGTCTATGCAGTGAAGGTGGAGGCAGAGCGGGGAA 360
OY	663 ACCGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCGGCCATCATGTGCGCAAGA 722
Db	361 ACCGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCGGCCATCATGTGCGCAAGA 420
OY	723 GAAGAACAAGCCAGATCTGTGGCCTCCGANGACGAAAAGGCTGACAGATMAATCAAGGC 782
Db	421 GAAGAACAAGCCAGATCTGTGGCCTCCGANGACGAAAAGGCTGACAGATMAATCAAGGC 480
OY	783 AGCAGAGAGGCCAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAGCTGAAGTATTCCAAT 842
Db	481 AGCAGAGAGGCCAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAGCTGAAGTATTCCAAT 540
OY	843 CCTGTGTCACACTCTGACACATAATGAGATGACAGACTTCACAGACTGTGGGCGA 902
Db	541 CCTGTGTCACACTCTGACACATAATGAGATGACAGACTTCACAGACTGTGGGCGA 600
OY	903 GCATATATGTACGCCGTTCCTCAACATGAGCCAAAGACTCCAAACACTATCTACTAGCC 959

Query Match	46.7%: Best Local Similarity	Score 555:	DB 10:	Length 702:
Matches 655:	Conservative	0:	Mismatches 2:	Indels 0:
			Gaps 0:	
QY	303	GGAAATGTCTATCAACGATGCGCTGAGCACTGCGCTGTGACTCTCGACATGTACTCTGCA	362	
Db	1	GGAAATGTCTATCAACGATGCGCTGAGCACTGCGCTGTGACTCTCGACATGTACTCTGCA	60	
QY	363	AATGATGAGATCCCTTTACCTGCGCATATATGACCCCTTACAAGGCAGCTACGGTGTGGA	422	
Db	61	AATGATGAGATCCCTTTACCTGCGCATATATGACCCCTTACAAGGCAGCTACGGTGTGGA	120	
QY	423	GGACCTGAGTATGCGCTGACCCAGCTAGCTCAAAACACATGAGATAGAGCTCGCGAA	482	
Db	121	GGACCTGAGTATGCGCTGACCCAGCTAGCTCAAAACACATGAGATAGAGCTCGCGAA	180	
QY	483	ACTCTCTNTGCAACAAGTCTTCCGGGAAACGGGAGTCCCTGAATGCCAGACTTGTGATGC	542	
Db	181	ACTCTCTNTGCAACAAGTCTTCCGGGAAACGGGAGTCCCTGAATGCCAGACTTGTGATGC	240	
QY	543	CATCAACCAACCTGCTGCTGCGGTATCCGCTGCTTCGTTATGATGATCAAGATAT	602	
Db	241	CATCAACCAACCTGCTGCTGCGGTATCCGCTGCTTCGTTATGATGATCAAGATAT	300	
QY	603	CCATGTGCGACCCCGGGGTGAAAGAGTCTATGACAGTGGAGGAGCGAGCGCGAA	662	
Db	301	CCATGTGCGACCCCGGGGTGAAAGAGTCTATGACAGTGGAGGAGCGAGCGCGAA	360	
QY	663	ACGGGCGACACTTTCTAAGTGTGAGGGGACCCGAGAGTCCGCCATCAATGTGGCGAAGG	722	
Db	361	ACGGGCGACACTTTCTAAGTGTGAGGGGACCCGAGAGTCCGCCATCAATGTGGCGAAGG	420	
QY	723	GAAAGAACAGGCGCCAGATCCTGCGCTCCGAGGCGAAAGGCTTAACAGATTAATTCAGGC	782	
Db	421	GAAAGAACAGGCGCCAGATCCTGCGCTCCGAGGCGCCAGGCTTAACAGTATTTCGAAT	480	
QY	783	AGCAGAGAGGCGCAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAGCTGAGCTATTTCGAAT	842	
Db	481	AGCAGAGAGGCGCAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAGCTGAGCTATTTCGAAT	540	
QY	843	CGTGGCTGCGACCTCTGACACACAACATTAATGAGATGCGAGACCTTCACTGACTGTGGCGGA	902	
Db	541	CGTGGCTGCGACCTCTGACACACAACATTAATGAGATGCGAGACCTTCACTGACTGTGGCGGA	600	
QY	903	GCACTATGTCAGGCCGTTTCTTCCAACATGGGCCAAGGACTCCAAACACTATTCCTACTGCGC	959	

Y 800 CAGTTGCGGAGGCGCAAGCTTAAGCTATTCGATTCCTGGCTCAGCTCTGA 859
|||||
Db 791 CAGTTGCGGAGGCGCAAGCTTAAGCTATTCGATTCCTGGCTCAGCTCTGA 850
Y 860 CACA 863
|||||
Db 851 CACA 854

RESULT 10
BI092184
LOCUS 796 bp mRNA linear EST 20-JUN-2001
DEFINITION 602856421F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4997837 5',
mRNA sequence.
ACCESSION BI092184
VERSION BI092184.1 GI:14510514
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM1025 row: i column: 06
High quality sequence stop: 793.
Location/Qualifiers
1. 796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4997837"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 206 a 214 c 230 g 146 t
ORIGIN

Query Match 47.88; Score 568; DB 10; Length 796;
Best Local Similarity 99.78; Pred. No. 2.5e-275;
Matches 668; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Y 352 GTAATCTGCAAAATGATGAGTCTTACCTGCGCATCATGAGCCCTTACCAAGCAGC 411
|||||
Db 42 GTAATCTGCAAAATGATGAGTCTTACCTGCGCATCATGAGCCCTTACCAAGCAGC 101
Y 412 TACGGTGTGAGAGACCTGAGTATCCCTACACCAAGTAGCTCAAAACATCATGATCA 471
|||||
Db 102 TACGGTGTGAGAGACCTGAGTATCCCTACACCAAGTAGCTCAAAACATCATGATCA 161
Y 472 GAGCTCGGCAAACTCTCTNTGGACAAGTCTTCGGGAAGGAGTCCCTGAATGCCAGC 531
|||||
Db 162 GAGCTCGGCAAACTCTCTNTGGACAAGTCTTCGGGAAGGAGTCCCTGAATGCCAGC 221
Y 532 ATTGTGATGCCATCAACCAAGCTGCTGACTGCTGGGTATCCGCTGCTTGTATGAG 591
|||||
Db 222 ATTGTGATGCCATCAACCAAGCTGCTGACTGCTGGGTATCCGCTGCTTGTATGAG 281
Y 592 ATCAAGGATATCATGTGCGCACCCCGGCTGAAGAGTCTATGACATGACAGTGGAGGCA 651

Db 282 ATCAAGGATATCATGTGCGCACCCCGGCTGAAGAGTCTATGACATGACAGTGGAGGCA 341
Y 652 GAGCGCGGAAACGGGCGACAGTCTAGAGTCTGAGGGGACCCGAGAGTCCGCAATCAAT 711
|||||
Db 342 GAGCGCGGAAACGGGCGACAGTCTAGAGTCTGAGGGGACCCGAGAGTCCGCAATCAAT 401
Y 712 GTGCGAGAGGGAAGAAACAGAGCCAGATCTGCGCTCCGAAGCAGAAAGAGCTGACAG 771
|||||
Db 402 GTGCGAGAGGGAAGAAACAGAGCCAGATCTGCGCTCCGAAGCAGAAAGAGCTGACAG 461
Y 772 ATAAATCAGCAGCAGAGAGAGCCAGTGTCTGCGGAAAGCCCAAGCTTAAGCTGAA 831
|||||
Db 462 ATAAATCAGCAGCAGAGAGAGCCAGTGTCTGCGGAAAGCCCAAGCTTAAGCTGAA 521
Y 832 GCTATTCGAATCTCTGCGTCTGACACATCAACATATGAGAGATGACAGCAGCTTCACTG 891
|||||
Db 522 GCTATTCGAATCTCTGCGTCTGACACATCAACATATGAGAGATGACAGCAGCTTCACTG 581
Y 892 ACTGTGCGCGAGCAGATGATGACGCGCTTCTCCAAACTGCGCAAGAGCTCCAACTATC 951
|||||
Db 582 ACTGTGCGCGAGCAGATGATGACGCGCTTCTCCAAACTGCGCAAGAGCTCCAACTATC 641
Y 952 CTACTGCCCTCCAACTGCGGATGTCACACAGATGTGCTCAGGCGCATGGTGTATAT 1011
|||||
Db 642 CTACTGCCCTCCAACTGCGGATGTCACACAGATGTGCTCAGGCGCATGGTGTATAT 701
Y 1012 GGAGCCCTCA 1021
|||||
Db 702 GGAGCCCTCA 711

RESULT 11
BE740453
LOCUS 660 bp mRNA linear EST 15-SEP-2000
DEFINITION 601595494F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949582 5',
mRNA sequence.
ACCESSION BE740453
VERSION BE740453.1 GI:10154445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DHP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM814 row: k column: 23
High quality sequence stop: 660.
Location/Qualifiers
1. 660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949582"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site:1; XhoI; Site:2;
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

Query Match 48.1%; Score 572; DB 10; Length 887;
 Best Local Similarity 99.7%; Pred. No. 2,5e-277;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

252 GAACATCCCATCCCTGTTAGACCGGATCGGATGTGTGACAGAGTCTCAGAGAAATTCG 311
 |||||||
 2 GAACATCCCATCCCTGTTAGACCGGATCGGATGTGTGACAGAGTCTCAGAGAAATTCG 61

312 CATCAACGTCCTGAGACAGTCCGCTGTGATCTCGACAAATGTAATCTCAATCATGATG 371
 |||||||
 62 CATCAACGTCCTGAGACAGTCCGCTGTGATCTCGACAAATGTAATCTCAATCATGATG 121

372 AGTCCTTTACCTGCGCATATGAGACCTTACAGGCAAGCTACGCTGTGAGAGACCTGCA 431
 |||||||
 122 AGTCCTTTACCTGCGCATATGAGACCTTACAGGCAAGCTACGCTGTGAGAGACCTGCA 181

432 GTATGCGGTCACCGATGCTCAAAACCAATGATGATGAGCTCGGCAACTCTCTMT 491
 |||||||
 182 GTATGCGGTCACCGATGCTCAAAACCAATGATGATGAGCTCGGCAACTCTCTMT 241

492 GGACAAAGTCTTCGCGGAGAGGAGTCCCTGATGCGACAGATTGTGATGCCATCAACCA 551
 |||||||
 242 GGACAAAGTCTTCGCGGAGAGGAGTCCCTGATGCGACAGATTGTGATGCCATCAACCA 301

552 AGCTCTGACTCTGCGGATGATCCCTGCTTACGATGATGATGATGATGATGATGATGATG 611
 |||||||
 302 AGCTCTGACTCTGCGGATGATCCCTGCTTACGATGATGATGATGATGATGATGATGATG 361

612 ACCCGGGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 671
 |||||||
 362 ACCCGGGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421

672 AGTTTACAGTCTGAGGAGGAGGAGAGTGGCATCAATGTGGCAGAGAGGAGAGAGACA 731
 |||||||
 422 AGTTTACAGTCTGAGGAGGAGGAGAGTGGCATCAATGTGGCAGAGAGGAGAGAGACA 481

732 GGGCCAGATCCCTGCGGCTCCGAGAGCAAGAAAGCTGAGACAGATTAATCAGGCGACAGAGA 791
 |||||||
 482 GGGCCAGATCCCTGCGGCTCCGAGAGCAAGAAAGCTGAGACAGATTAATCAGGCGACAGAGA 541

792 GGGCCAGATCCCTGCGGCTCCGAGAGCAAGAAAGCTGAGACAGATTAATCAGGCGACAGAGA 851
 |||||||
 542 GGGCCAGATCCCTGCGGCTCCGAGAGCAAGAAAGCTGAGACAGATTAATCAGGCGACAGAGA 601

852 AGCTTGACACAAATATGAGATGACAGACAGCTTCACTGATGTGGCCGACAGATATGT 911
 |||||||
 602 AGCTTGACACAAATATGAGATGACAGACAGCTTCACTGATGTGGCCGACAGATATGT 661

912 CAGCGGCTTCCCA 925
 |||||||
 662 CAGCGGCTTCCCA 675

RESULT 9
 BM476304 1110 bp mRNA linear EST 05-FEB-2002
 LOCUS BM476304
 DEFINITION AGNCOURT_6479118 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5558923
 5', mRNA sequence.
 ACCESSION BM476304
 VERSION BM476304.1 GI:18525346
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1282 row: k column: 20
 High quality sequence start: 8
 High quality sequence stop: 678.
 Location/Qualifiers
 1. 1110
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5558923"
 /clone_lib="NIH_MGC_88"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: small intestine; Vector: pcMV-SpO6;
 Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
 oligo-dr primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 249 a 322 c 326 g 213 t

Query Match 48.1%; Score 571; DB 10; Length 1110;
 Best Local Similarity 99.6%; Pred. No. 8.4e-277;
 Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

140 GCCCGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 199
 |||||||
 131 GCCCGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 190

200 CCTGGGTGTGAGGCAATGAGGCGGATTCACCGGATCTGAGGCTGGTTGAACATCC 259
 |||||||
 191 CCTGGGTGTGAGGCAATGAGGCGGATTCACCGGATCTGAGGCTGGTTGAACATCC 250

260 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGAAATGTCTATCAACG 319
 |||||||
 251 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGAAATGTCTATCAACG 310

320 TGCCTGAGAGTCTGCTGATCTCTGACATGTAATCTGCAATTCATGAGAGTCTTT 379
 |||||||
 311 TGCCTGAGAGTCTGCTGATCTCTGACATGTAATCTGCAATTCATGAGAGTCTTT 370

380 ACCTGGCATCATGAGACCTTACAAAGCAAGCTAGAGTGTGAGAGACCTGAGATGCCG 439
 |||||||
 371 ACCTGGCATCATGAGACCTTACAAAGCAAGCTAGAGTGTGAGAGACCTGAGATGCCG 430

440 TCACCCAGCTAGCTCAAAACCAACATGATGATGATGATGATGATGATGATGATGATGATG 499
 |||||||
 431 TCACCCAGCTAGCTCAAAACCAACATGATGATGATGATGATGATGATGATGATGATGATG 490

500 TCTTCCGGGAAGGAGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 559
 |||||||
 491 TCTTCCGGGAAGGAGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 550

560 ACTGCTGGGATGATGCGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATG 619
 |||||||
 551 ACTGCTGGGATGATGCGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATG 610

620 TGAAGAAGTCTATGAGATGAGTGGAGAGAGCGGCGGGAAGAGGCGGACAGATTCTAG 679
 |||||||
 611 TGAAGAAGTCTATGAGATGAGTGGAGAGAGCGGCGGGAAGAGGCGGACAGATTCTAG 670

680 AGTCTGAGGAGACCGGAGAGTGGCCATCATATGTGGCAGAGAGGAGAGAGAGGCGGCGAG 739
 |||||||
 671 AGTCTGAGGAGACCGGAGAGTGGCCATCATATGTGGCAGAGAGGAGAGAGAGGCGGCGAG 730

740 TCTTGCCCTCCGAGACAGAAAGGCTGGAACAGATTAATCAGCAGAGAGAGAGGCGGCGAG 799
 |||||||
 731 TCTTGCCCTCCGAGACAGAAAGGCTGGAACAGATTAATCAGCAGAGAGAGAGGCGGCGAG 790

Db 661 TCCAGGCTACAGATGCAAGT 680

RESULT 7

LOCUS BM461614 1011 bp mRNA linear EST 05-FEB-2002

DEFINITION AGENCOURT_6423435 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5501391

ACCESSION BM461614

VERSION 5', mRNA sequence.

KEYWORDS BM461614.1 GI:18510654

SOURCE EST.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1011)

NIH-MGC http://mhc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

plate: LLM12137 row: n column: 16

High quality sequence stop: 622.

Location/Qualifiers

1. 1011

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5501391"

/clone_lib="NIH_MGC_67"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo Average insert size 1.75 kb. Library constructed by Life Technologies."

Technologies."

BASE COUNT 232 a 289 c 288 g 201 t 1 others

ORIGIN

Query Match 48 28; Score 573; DB 10; Length 1011;

Best Local Similarity 99.78; Pred. No. 8e-278;

Matches 673; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCCGCCCTCTGTGATTGCCGGAACACCGTGTACTGTCTGCGCGAGCAGAGG 199

Db 61 GCCCGCCCTCTGTGATTGCCGGAACACCGTGTACTGTCTGCGCGAGCAGAGG 120

Qy 200 CTTGGGTGTGAGCGAATGGCGGATTCCACCGGATCCTGGAGCCTGTGGAAATCC 259

Db 121 CTTGGGTGTGAGCGAATGGCGGATTCCACCGGATCCTGGAGCCTGTGGAAATCC 180

Qy 260 TCATCCCTGTGTAGACCGGATTCGATGTGCGAGAGTCTCAAGGAATGTGCATCAAG 319

Db 181 TCATCCCTGTGTAGACCGGATTCGATGTGCGAGAGTCTCAAGGAATGTGCATCAAG 240

Qy 320 TGCCGTGACAGTGGCTGTGACTCTGACAAATGTAACTCTGCAATCGATGGAGTCTTT 379

Db 241 TGCCGTGACAGTGGCTGTGACTCTGACAAATGTAACTCTGCAATCGATGGAGTCTTT 300

Qy 380 ACCTGCGATCATGAGACCTTACAAAGCAAGTACGGTGTGAGAGACCTTGAGCTTGGCG 439

Db 301 ACCTGCGATCATGAGACCTTACAAAGCAAGTACGGTGTGAGAGACCTTGAGCTTGGCG 360

Qy 440 TCACCGAGCTAGCTCAAAACCAATGAGATCGAGCTCGGCAAACTCTCTNTGCAAAAG 499

Db 361 TCACCGAGCTAGCTCAAAACCAATGAGATCGAGCTCGGCAAACTCTCTNTGCAAAAG 420

Qy 500 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGTATGCCATCAACCAAGTCTG 559

Db 421 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGTATGCCATCAACCAAGTCTG 480

Qy 560 ACTGCTGGGATATCCGCTGCTTGTATGATGATCAAGATATCCATGATCCACCCGGG 619

Db 481 ACTGCTGGGATATCCGCTGCTTGTATGATGATCAAGATATCCATGATCCACCCGGG 540

Qy 620 TGAAGAGTCTATGACATGATGAGTGGAGGAGCGCGGAAACGGGCGCAAGTCTTAG 679

Db 541 TGAAGAGTCTATGACATGATGAGTGGAGGAGCGCGGAAACGGGCGCAAGTCTTAG 600

Qy 680 AGTGTGAGGGGACCCGAGAGTGGCCATCAATGTGCGAAGGAGCAAAACAGGCCAGA 739

Db 601 AGTGTGAGGGGACCCGAGAGTGGCCATCAATGTGCGAAGGAGCAAAACAGGCCAGA 660

Qy 740 TCTTGCCCTCCGAGACGAAAGGCTGAACGATTAATTCAGGCGAGGAGGCGAGTG 799

Db 661 TCTTGCCCTCCGAGACGAAAGGCTGAACGATTAATTCAGGCGAGGAGGCGAGTG 720

Qy 800 CAGTTCTGGCGAAG 814

Db 721 CAGTTCTGGCGAAG 735

RESULT 8

LOCUS BM045426 887 bp mRNA linear EST 07-NOV-2001

DEFINITION 60362354F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5449259 5', mRNA sequence.

ACCESSION BM045426

VERSION BM045426.1 GI:16774693

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 887)

NIH-MGC http://mhc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

plate: LLM1936 row: b column: 12

High quality sequence stop: 768.

Location/Qualifiers

1. 887

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5449259"

/clone_lib="NIH_MGC_40"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Note: This is a NIH_MGC Library."

BASE COUNT 237 a 232 c 247 g 171 t

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1,8e-295;
Matches 708; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 203 GGGTGGTGGAGGCAATGGGCGGATTCACCGGATTCGAGCCTGGTTGAAACATCTCA 262
|||||
Db 1 GGGTGGTGGAGGCAATGGGCGGATTCACCGGATTCGAGCCTGGTTGAAACATCTCA 60
OY 263 TCCCTGTGTAGACCGCGATTCGATATGTGACAGAGTCTCAAGAAATTTGTCATACGTCG 322
|||||
Db 61 TCCCTGTGTAGACCGCGATTCGATATGTGACAGAGTCTCAAGAAATTTGTCATACGTCG 120
OY 323 CTGAGCAGTCCGCTGTGACTCTGACATATGTGATCTGCAAAATCGATGAGTCCCTTACC 382
|||||
Db 121 CTGAGCAGTCCGCTGTGACTCTGACATATGTGATCTGCAAAATCGATGAGTCCCTTACC 180
OY 383 TCGCATCATGAGACCTTACAAAGCAAGCTACGGTGTGGAGACCTGATATGCGTCA 442
|||||
Db 181 TCGCATCATGAGACCTTACAAAGCAAGCTACGGTGTGGAGACCTGATATGCGTCA 240
OY 443 CCCAGCTAGCTCAAAACACCATGATGATGAGCTGGCAAACTCTGTTGACAAAGTCT 502
|||||
Db 241 CCCAGCTAGCTCAAAACACCATGATGATGAGCTGGCAAACTCTGTTGACAAAGTCT 300
OY 503 TCCGGGAGGAGGAGTCCGTAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTGACT 562
|||||
Db 301 TCCGGGAGGAGGAGTCCGTAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTGACT 360
OY 563 GGTGGGATTCGCGTCCCTGCTTATGATGATCAAGATATCCATGTGCCACCCGGGTGA 622
|||||
Db 361 GGTGGGATTCGCGTCCCTGCTTATGATGATCAAGATATCCATGTGCCACCCGGGTGA 420
OY 623 AAGAGTCTATGCAAGATGACAGGTGAGGACAGAGCGGCAAGGCGGACAGTTCTAGAGT 682
|||||
Db 421 AAGAGTCTATGCAAGATGACAGGTGAGGACAGAGCGGCAAGGCGGACAGTTCTAGAGT 480
OY 683 CTGAGGAGGAGGAGGAGTGGGCGATCAATGTGGCAGAAAGGAGAAACAGGCCCAATCC 742
|||||
Db 481 CTGAGGAGGAGGAGGAGTGGGCGATCAATGTGGCAGAAAGGAGAAACAGGCCCAATCC 540
OY 743 TGGCCTCCGAGACAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGAGGCGAGTGCAG 802
|||||
Db 541 TGGCCTCCGAGACAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGAGGCGAGTGCAG 600
OY 803 TTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTCGAAATCTGCTGACGCTGTGACAC 862
|||||
Db 601 TTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTCGAAATCTGCTGACGCTGTGACAC 660
OY 863 AACATTAAGAGATGACAGGCTTCACTGACTGTGGCGAGAGCTATGTC 912
|||||
Db 661 AACATTAAGAGATGACAGGCTTCACTGACTGTGGCGAGAGCTATGTC 710

RESULT 6
BM456748 1053 bp mRNA linear EST 05-FEB-2002
LOCUS BM456748
DEFINITION AGENCOURT_6403923 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583626
5', mRNA sequence.
ACCESSION BM456748
VERSION BM456748.1 GI:18505788
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML12347 row: a column: 03
High quality sequence stop: 621.
Location/Qualifiers

FEATURES
source
1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5583626"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site: 1; Noti:
Site: 2; SalI: Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by life technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 269 a 285 c 290 g 209 t
ORIGIN

Query Match 48.7%; Score 578; DB 10; Length 1053;
Best Local Similarity 99.7%; Pred. No. 2,4e-280;
Matches 678; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 419 TGGAGAGACCCCTGGTATGCGGTCACCAGCTAGCTCAAAACCAATGAGATGACCTCG 478
|||||
Db 1 TGGAGAGACCCCTGGTATGCGGTCACCAGCTAGCTCAAAACCAATGAGATGACCTCG 60
OY 479 GCAAACTCTCTNTGACAAAGCTTCCGGGAACGGGAGTCCCTGAATGCGCAGATTGTGG 538
|||||
Db 61 GCAAACTCTCTNTGACAAAGCTTCCGGGAACGGGAGTCCCTGAATGCGCAGATTGTGG 120
OY 539 ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATTCGCTGCTTCGTTATGATCAAG 598
|||||
Db 121 ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATTCGCTGCTTCGTTATGATCAAG 180
OY 599 ATATCATGTCGCCACCCCGGGTGAAGAGTCTATGAGATGACAGGAGGAGGCGAGACGGC 658
|||||
Db 181 ATATCATGTCGCCACCCCGGGTGAAGAGTCTATGAGATGACAGGAGGAGGCGAGACGGC 240
OY 659 GGAAGAGGCGCAGAGTCTAGAGTCTGAGGGAGCCGAGAGTGGCCATCAATGTGGCAG 718
|||||
Db 241 GGAAGAGGCGCAGAGTCTAGAGTCTGAGGGAGCCGAGAGTGGCCATCAATGTGGCAG 300
OY 719 AAGGAGAAAGAACAGGCGCCAGATCTGGCTCCGAGCAGAAAGGCTGAACAGATTAATC 778
|||||
Db 301 AAGGAGAAAGAACAGGCGCCAGATCTGGCTCCGAGCAGAAAGGCTGAACAGATTAATC 360
OY 779 AGCAGCAGAGAGGAGGCGAGTCTGAGTCTGGCGAAGGCCAAAGCTTAAGCTGAAGCTATTTC 838
|||||
Db 361 AGCAGCAGAGAGGAGGCGAGTCTGAGTCTGGCGAAGGCCAAAGCTTAAGCTGAAGCTATTTC 420
OY 839 GAATCTGCTGCTGAGCTGACACAAATTAATGAGATGACAGCAGCTTCACTGACTGTGG 898
|||||
Db 421 GAATCTGCTGCTGAGCTGACACAAATTAATGAGATGACAGCAGCTTCACTGACTGTGG 480
OY 899 CCGAGCAGTATGTCAGCGCGTTCTCCAAATGGCCAGAGACTCAACACTATCTTCTGTC 958
|||||
Db 481 CCGAGCAGTATGTCAGCGCGTTCTCCAAATGGCCAGAGACTCAACACTATCTTCTGTC 540
OY 959 CCGCACAACCTGGCGATGTCACAGCATGATGGCTGAGGCAATGGGTATATGAGAGGCC 1018
|||||
Db 541 CCGCACAACCTGGCGATGTCACAGCATGATGGCTGAGGCAATGGGTATATGAGAGGCC 600
OY 1019 TCACCAAGCCCGCAGTGGCAGGAGCTCCAGACTCTTCCAGTGGAGCAGCAGAGATG 1078
|||||
Db 601 TCACCAAGCCCGCAGTGGCAGGAGCTCCAGACTCTTCCAGTGGAGCAGCAGAGATG 660
OY 1079 TCCAGGCTACAGATGCAAGT 1098
|||||

BM453159 1061 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6387825 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5529785
DEFINITION 5', mRNA sequence.
ACCESSION BM453159
VERSION BM453159.1 GI:18502199
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM12208 row: m column: 18
High quality sequence stop: 661.
Location/Qualifiers
1..1061
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5529785"
/clone_lib="NIH_MGC_71"
/tissue_type="telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site:1; Noti:
Site:2; Salt: Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb.
BASE COUNT 259 a 298 c 279 g 225 t
ORIGIN
Query Match 54.9%; Score 652; DB 10; Length 1061;
Best Local Similarity 99.7%; Pred. No. 0; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 2;
Oy 345 CGACATGTAAGTCTGCAATGATGAGTCTTCTTACTGCGCATGAGCCCTTACAA 404
Db 1 CGACATGTAAGTCTGCAATGATGAGTCTTCTTACTGCGCATGAGCCCTTACAA 60
Oy 405 GGCAGACTAGGCTGTGAGGAGCCCTGATGATGCGCTGACCCAGCTCAACAAACAT 464
Db 61 GGCAGACTAGGCTGTGAGGAGCCCTGATGATGCGCTGACCCAGCTCAACAAACAT 120
Oy 465 GAGATCAGAGCTCGCAAACTCTCTNTGCAAAAGTCTCCGGGAGCGGAGTCCCTGAA 524
Db 121 GAGATCAGAGCTCGCAAACTCTCTNTGCAAAAGTCTCCGGGAGCGGAGTCCCTGAA 180
Oy 525 TGCCACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 584
Db 181 TGCCACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Oy 585 TTATGAGATCAAGATATATCATGTCACCCCGGATGAAAGATCTATGACATGACAGT 644
Db 241 TTATGAGATCAAGATATATCATGTCACCCCGGATGAAAGATCTATGACATGACAGT 300
Oy 645 GGAGGAGAGAGCGGGGAG 704
Db 301 GGAGGAGAGAGCGGGGAG 360
Oy 705 CATCATGTGTGGCAG 764
Db 361 CATCATGTGTGGCAG 420

Oy 765 TGACAGATTAATTCAGGAG 824
Db 421 TGACAGATTAATTCAGGAG 480
Oy 825 AGCTAAGCTATTCAGATTCCTGGCTGAGCTGACACAACTAATGAGATGAGAGAGAG 884
Db 481 AGCTAAGCTATTCAGATTCCTGGCTGAGCTGACACAACTAATGAGATGAGAGAGAG 540
Oy 885 TTCAGTACTGTGGCGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 944
Db 541 TTCAGTACTGTGGCGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 600
Oy 945 CACTATCTACTGCGCTCCACACCTGGGATGATGATGATGATGATGATGATGATGATGATG 1004
Db 601 CACTATCTACTGCGCTCCACACCTGGGATGATGATGATGATGATGATGATGATGATGATG 660
Oy 1005 TTATATGAGAGCGCTCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1064
Db 661 TTATATGAGAGCGCTCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Oy 1065 GAGCAGACAGATGTCCAGAGGTACAGATGCAAGT 1098
Db 721 GAGCAGACAGATGTCCAGAGGTACAGATGCAAGT 754

RESULT 5
BM423347 1095 bp mRNA linear EST 29-JAN-2002
LOCUS AGENCOURT 6402181 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5516467
DEFINITION 5', mRNA sequence.
ACCESSION BM423347
VERSION BM423347.1 GI:18391559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1095)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM2017 row: b column: 20
High quality sequence stop: 503.
BASE COUNT 285 a 282 c 302 g 226 t
ORIGIN
Query Match 51.2%; Score 608; DB 10; Length 1095;
FEATURES
source
1..1095
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5516467"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1; XhoI; Site:2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

Oy 429 TGAGTATGCCCTCACCCAGCTAGCTCAACACCATGATCAGAGCTCGGCAAACTCTC 488
Db 241 TGAGTATGCCCTCACCCAGCTAGCTCAACACCATGATCAGAGCTCGGCAAACTCTC 300
Oy 489 TTMGACAAAGTCTTCCGGGAAGGGAGTCCCTGAATGCCAGCTTTGGATGCCATCAA 548
Db 301 TCTGACAAAGTCTTCCGGGAAGGGAGTCCCTGAATGCCAGCTTTGGATGCCATCAA 360
Oy 549 CCAAGCTGCTGATGCTGCGGGTATCCGCTGCTTCGCTTATGATGATCAAGATATCATGT 608
Db 361 CCAAGCTGCTGATGCTGCGGGTATCCGCTGCTTCGCTTATGATGATCAAGATATCATGT 420
Oy 609 GCCACCCCGGGGTGAAGAGTCTATGACATGACAGTGGAGCGAGCGCGGAAACGGGC 668
Db 421 GCCACCCCGGGGTGAAGAGTCTATGACATGACAGTGGAGCGAGCGCGGAAACGGGC 480
Oy 669 CACAGTCTAGAGTCTGAGGGGACCGGAGTGGCCATCAATGTGGCAGAGGAGAA 728
Db 481 CACAGTCTAGAGTCTGAGGGGACCGGAGTGGCCATCAATGTGGCAGAGGAGAA 540
Oy 729 ACAGGCCAGATCCTGGCTCCGAGCAGAAAAGGCTGAACAGATTAATCAGCAGCAGG 788
Db 541 ACAGGCCAGATCCTGGCTCCGAGCAGAAAAGGCTGAACAGATTAATCAGCAGCAGG 600
Oy 789 AGAGGCCAGTGCAGTCTGGCGGAAGCCAAAGCTAAAGCTGAAGCTATTGCAATCTGGC 848
Db 601 AGAGGCCAGTGCAGTCTGGCGGAAGCCAAAGCTAAAGCTATTGCAATCTGGC 660
Oy 849 TGCAGCTCTGACACAACTAATGAGATGACAGCTTCACTGACTGGCGCGGAGTA 908
Db 661 TGCAGCTCTGACACAACTAATGAGATGACAGCTTCACTGACTGGCGCGGAGTA 720
Oy 909 TGTACAGGCGTCTCCAACTGGCCAGAGACTCCACACTATCTACTGCGCTTCAACCC 968
Db 721 TGTACAGGCGTCTCCAACTGGCCAGAGACTCCACACTATCTACTGCGCTTCAACCC 780
Oy 969 TGGCGATGTACACAGCATGGTGGCTCAGGCCAATGGGTATATGAGAGCCCTCACCAAGC 1028
Db 781 TGGCGATGTACACAGCATGGTGGCTCAGGCCAATGGGTATATGAGAGCCCTCACCAAGC 840
Oy 1029 CCCAGT 1034
Db 841 CCCAGT 846

RESULT 3
AL525168 849 bp mRNA linear EST 13-FEB-2001
LOCUS AL525168 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC005YF20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL525168
VERSION AL525168.1 GI:12788661
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 849)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 849
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC005YF20"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
```

```

/lab_host="DHL08"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 195 a 228 c 260 g 165 t 1 others
ORIGIN

Query Match 55.6% Score 660; DB 9; Length 849;
Best Local Similarity 99.7% Pred. No. 0;
Matches 760; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 140 GCCCGGCTCTCTGATGTTGCCGAAACACCGTGAATCTGTCGCGCAGCAGAG 199
Db 88 GCCCGGCTCTCTGATGTTGCCGAAACACCGTGAATCTGTCGCGCAGCAGAG 147
Oy 200 CTTGGTGTGAGCGAATGGGCGGATTCACCGGATCTTGAGCCTGTGTAATCC 259
Db 148 CTTGGTGTGAGCGAATGGGCGGATTCACCGGATCTTGAGCCTGTGTAATCC 207
Oy 260 TCATCCCTGTGTAGCCGGATGCGATATGACAGAGCTCAAGAAATGTCATCAAG 319
Db 208 TCATCCCTGTGTAGCCGGATGCGATATGACAGAGCTCAAGAAATGTCATCAAG 267
Oy 320 TGCCTGAGCACTGCGCTGTGACTCTGACATGTACTGTGCAATGATGATGACTCTT 379
Db 268 TGCCTGAGCACTGCGCTGTGACTCTGACATGTACTGTGCAATGATGATGACTCTT 327
Oy 380 ACTTGGCATATGAGACCTTACAGGCAAGCTACGTTGTGAGAGACCTTGATGCGC 439
Db 328 ACTTGGCATATGAGACCTTACAGGCAAGCTACGTTGTGAGAGACCTTGATGCGC 387
Oy 440 TCACCGAGCTAGCTCAACAAACCATGAGATGAGAGCTCGGCAACTCTCTTGAACAAG 499
Db 388 TCACCGAGCTAGCTCAACAAACCATGAGATGAGAGCTCGGCAACTCTCTTGAACAAG 447
Oy 500 TCTTCCGGGAACCGGACTCCCTGATGCCAGCATTTGTGATGTCATCAACCACTGCTG 559
Db 448 TCTTCCGGGAACCGGACTCCCTGATGCCAGCATTTGTGATGTCATCAACCACTGCTG 507
Oy 560 ACTGCTGGGATATCCGCTGCTTCGTTATGATGATCAAGATATCCATGTGCCACCCGGC 619
Db 508 ACTGCTGGGATATCCGCTGCTTCGTTATGATGATCAAGATATCCATGTGCCACCCGGC 567
Oy 620 TGAAGAGTCTATGACATGAGTGTGAGGAGCGGCGGAAACGGGCCCACTTTAG 679
Db 568 TGAAGAGTCTATGACATGAGTGTGAGGAGCGGCGGAAACGGGCCCACTTTAG 627
Oy 680 AGTGTAGGGGACCCGAGAGTCGCGCATATGTGGCAGAGGAAAGAACAGGCCAGCA 739
Db 628 AGTGTAGGGGACCCGAGAGTCGCGCATATGTGGCAGAGGAAAGAACAGGCCAGCA 687
Oy 740 TCTTGGCTCTCGAAGCAGAAAGGCTCAACAGATTAATCAGGCGCAGAGAGCCAGTG 799
Db 688 TCTTGGCTCTCGAAGCAGAAAGGCTCAACAGATTAATCAGGCGCAGAGAGCCAGTG 747
Oy 800 CAGTTTGGGGAAGGCCAAGGCTAAAGCTAAGTATTCGAATCTGCTGACACTCTGA 859
Db 748 CAGTTTGGGGAAGGCCAAGGCTAAAGCTAAGTATTCGAATCTGCTGACACTCTGA 807
Oy 860 CACAACATAATGAGATGACAGACTTCACTGACTGTGGCG 901
Db 808 CACAACATAATGAGATGACAGACTTCACTGACTGTGGCG 849

RESULT 4
```

```

CDS
18. 1046
/codon_start=1
/product="HSPC108"
/protein_id="AF29073.1"
/db_xref="GI:5841440"
/translation="MRGSLASGRACRASSCLPNTVYLVFPOEAMVVERKGRHRI
LEPDLNLIPLVDRIYVOSLKEIVINVEQSAVTLNDVLTQIDGVLYLRIMDPYKAS
YEDDEPRAVTLQOTMRSELKSLDKYFERESLVAISVDAINQADWGRICLR
YEKDIHPVPRVESKOMQVABERRKRTYVESGRTSAINVEAGKKQAOILASEAR
KAEQINQAGASAVLAKAKAKAEIRLIAALQHNQNDAAISLTVARQVYASFKLA
KDSNTILPSPNPBDVTSVNAQMGVYALTRAPVPGTPDLSLSSSSRVOGTDSLDE
ELDRYKMS"

BASE COUNT      313 a      310 c      337 g      258 t
ORIGIN
Query Match      63.6%: Score 755; DB 11; Length 1218;
Best Local Similarity 99.6%: Pred. No. 0;
Matches 955; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 140 GCCGGCCCTCCTGGATGCCCCGAAACACCGGTGACTGTTGCGCGCAGCAGAGG 199
    |||||
Db 52 GCCGGCCCTCCTGGATGCCCCGAAACACCGGTGACTGTTGCGCGCAGCAGAGG 111
    |||||

Oy 200 CCTGGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGTTGAACATCC 259
    |||||
Db 112 CCTGGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGTTGAACATCC 171
    |||||

Oy 260 TCATCCCTGTGTAGACCGGATCCCATATGTGCAGAGTCTCAAGGAAATGTCAACAAG 319
    |||||
Db 172 TCATCCCTGTGTAGACCGGATCCCATATGTGCAGAGTCTCAAGGAAATGTCAACAAG 231
    |||||

Oy 320 TGCCTGAGCAGTGGGTGTGATCTGTGACATGTACTGTGCAAAATGATGAGTCTTT 379
    |||||
Db 232 TGCCTGAGCAGTGGGTGTGATCTGTGACATGTACTGTGCAAAATGATGAGTCTTT 291
    |||||

Oy 380 ACCTGGCATCATGACACCTTACAAAGGCAAGCTACGGTGTGGAGACCTGAGTATGCCG 439
    |||||
Db 292 ACCTGGCATCATGACACCTTACAAAGGCAAGCTACGGTGTGGAGACCTGAGTATGCCG 351
    |||||

Oy 440 TCACCAGCTAGCTCAAAACCATGATGATGAGTCCGCAAACTCTGTTGACACAAG 499
    |||||
Db 352 TCACCAGCTAGCTCAAAACCATGATGATGAGTCCGCAAACTCTCTGTCGACACAAG 411
    |||||

Oy 500 TCTTCCGGAGAGGAGTCCCTGTAATGCCAGCATTTGTGATGTCATCAACCAAGTCTG 559
    |||||
Db 412 TCTTCCGGAGAGGAGTCCCTGTAATGCCAGCATTTGTGATGTCATCAACCAAGTCTG 471
    |||||

Oy 560 ACTGCTGGGGGTATCCGCTGCTTCGTTATGATGATCAAGGATATTCATGTCACCCCGG 619
    |||||
Db 472 ACTGCTGGGGGTATCCGCTGCTTCGTTATGATGATCAAGGATATTCATGTCACCCCGG 531
    |||||

Oy 620 TGAAGAGTCTATGACAGATGACAGTGGAGGAGCGCGGAAAGGAGCCACAGTTCTAG 679
    |||||
Db 532 TGAAGAGTCTATGACAGATGACAGTGGAGGAGCGCGGAAAGGAGCCACAGTTCTAG 591
    |||||

Oy 680 AGTCTGAGGGAGCCGAGAGTCCGATCAATGTGCGAAGAGGAAAGCAAGGCCGAGA 739
    |||||
Db 592 AGTCTGAGGGAGCCGAGAGTCCGATCAATGTGCGAAGAGGAAAGCAAGGCCGAGA 651
    |||||

Oy 740 TCTGTGCTCCGAGAGGAAAGGCTGACAGATTAATTCAGGAGCAGAGAGGCCGAGT 799
    |||||
Db 652 TCTGTGCTCCGAGAGGAAAGGCTGACAGATTAATTCAGGAGCAGAGAGGCCGAGT 711
    |||||

Oy 800 CAGTTTGGCGAAGCGCAAGGCTAAAGCTATTCGAATCTTGAGTGCAGCTCTGA 859
    |||||
Db 712 CAGTTTGGCGAAGCGCAAGGCTAAAGCTATTCGAATCTTGAGTGCAGCTCTGA 771
    |||||

Oy 860 CACAACATAATGAGATGACAGCAGCTTCACTGACTGTGCCGAGCACTATGTCAAGCCGT 919
    |||||
Db 772 CACAACATAATGAGATGACAGCAGCTTCACTGACTGTGCCGAGCACTATGTCAAGCCGT 831
    |||||

Oy 920 TCTCCAAACTGGCGAAGAGACTCCACACACTATCTACTGCCCTCCAAACCTGGCGATGTCA 979
    |||||

```

```

Db 832 TCTCCAACTGGCGAAGAGACTCCACACTATCTCTGCTCCCAACCTGGCGATGTCA 891
    |||||

Oy 980 CCAGCATGTGCTCAGGCGCATGAGTGTATGAGAGCCCTCACCAAGGCCAGTGCAG 1039
    |||||
Db 892 CCAGCATGTGCTCAGGCGCATGAGTGTATGAGAGCCCTCACCAAGGCCAGTGCAG 951
    |||||

Oy 1040 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACATGCAAGT 1098
    |||||
Db 952 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACATGCAAGT 1010
    |||||

RESULT 2
AL541792      935 bp      mRNA      linear      EST 16-FEB-2001
LOCUS        AL541792.LTI_FL002.PL1 Homo sapiens cDNA clone CS0DE007YE18 5 prime
DEFINITION   , mRNA sequence.
ACCESSION    AL541792.1 GI:12873201
VERSION      AL541792.1
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 935)
              Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
              Full-length cDNA libraries and normalization
              Unpublished (2001)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="LTI_FL002.PL1"
/lab_host="DH10B"
/note="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a Notti-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9600 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
liang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      239 a      246 c      265 g      181 t      4 others
ORIGIN
Query Match      58.3%: Score 693; DB 9; Length 935;
Best Local Similarity 99.6%: Pred. No. 0;
Matches 843; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 189 GCACGAGAGGCTGGGTGTGAGAGCGAATGGCCGATTCACCGGATCTCGAGCTCG 248
    |||||
Db 1 GCACGAGAGGCTGGGTGTGAGAGCGAATGGCCGATTCACCGGATCTCGAGCTCG 60
    |||||

Oy 249 TTTGACATCCTCATCCCTGTGTAGACGGATCCGATGTGTCAGAGTCTCAAGGAAT 308
    |||||
Db 61 TTTGACATCCTCATCCCTGTGTAGACGGATCCGATGTGTCAGAGTCTCAAGGAAT 120
    |||||

Oy 309 TGTATCAACGTCGCTGAGCAGTCCGCTGTGACTCTTGACATGTAACTGTCAAAATCGA 368
    |||||
Db 121 TGTATCAACGTCGCTGAGCAGTCCGCTGTGACTCTTGACATGTAACTGTCAAAATCGA 180
    |||||

Oy 369 TGGAGTCTTTTACCTGGCGCATGATGACACCTTACAAAGGCAAGCTAGCGTGGAGGCC 428
    |||||
Db 181 TGGAGTCTTTTACCTGGCGCATGATGACACCTTACAAAGGCAAGCTAGCGTGGAGGCC 240
    |||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:22:53 : Search time 1701.18 Seconds
(without alignments)
9425.447 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GCGTTCGAGACNACCGCT.....GGAACGAGATTTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755	63.6	1218	11	AF161458 Homo sapi
2	693	58.3	935	9	AL541792 AL541792
3	660	55.6	849	9	AL525168 AL525168
4	652	54.9	1061	10	BM453159 BM453159
5	608	51.2	1095	10	BM423347 BM423347
6	578	48.7	1053	10	BM456748 BM456748
7	573	48.2	1011	10	BM461614 BM461614
8	572	48.1	887	10	BM045426 BM045426
9	571	48.1	1110	10	BM476304 BM476304
10	568	47.8	796	10	BI092184 BI092184
11	558	47.0	660	10	BE740453 BE740453
12	555	46.7	702	10	BI838402 BI838402
13	552	46.5	749	10	BC036463 BC036463
14	539	45.4	770	10	BC828070 BC828070
15	536	45.1	1037	10	BM471641 BM471641
16	534	44.9	853	10	BE314702 BE314702
17	531	44.7	820	10	BG685336 BG685336

18	518	43.6	994	10	BM468527 BM468527
19	516	43.4	918	10	BM459766 BM459766
20	515	43.4	644	10	BE273795 BE273795
21	513	43.2	616	10	BM045444 BM045444
22	504	42.4	614	10	BE256589 BE256589
23	504	42.4	688	10	BE684435 BE684435
24	501	42.2	728	10	BI522493 BI522493
25	500	42.1	689	10	BM015578 BM015578
26	500	42.1	920	10	BI862105 BI862105
27	498	41.9	962	10	BG112719 BG112719
28	487	41.0	998	10	BM468407 BM468407
29	483	40.7	737	10	BG700808 BG700808
30	483	40.7	841	10	BE904561 BE904561
31	482	40.6	739	10	BG574270 BG574270
32	478	40.2	934	10	BF344401 BF344401
33	477	39.7	730	10	BI761597 BI761597
34	472	39.7	738	10	BG767007 BG767007
35	471	39.6	778	10	BG681190 BG681190
36	470	39.6	752	10	BI552983 BI552983
37	469	39.5	785	10	BM009759 BM009759
38	464	39.1	836	10	BE734242 BE734242
39	463	39.0	691	10	BE297150 BE297150
40	463	39.0	683	10	BF311329 BF311329
41	463	39.0	707	10	BI533283 BI533283
42	463	39.0	725	10	BE275386 BE275386
43	463	39.0	736	10	BI198047 BI198047
44	463	39.0	845	10	BE893277 BE893277
45	463	39.0	871	10	BE794548 BE794548

ALIGNMENTS

RESULT 1	AF161458	1218 bp	MRNA	linear	HTC 22-MAY-2001
LOCUS	AF161458				
DEFINITION	Homo sapiens HSPC108 mRNA, complete cds.				
ACCESSION	AF161458				
VERSION	AF161458.1	GI:6841439			
KEYWORDS	HTC.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Zhang,Q.H., Ye,M., Wu,X.Y., Ren,S.X., Zhao,M., Zhao,C.J., Fu,G., Shen,Y., Fan,H.Y., Lu,G., Zhong,M., Xu,X.R., Han,Z.G., Zhang,J.W., Tao,J., Huang,Q.H., Zhou,J., Hu,G.X., Gu,J., Chen,S.J., and Chen,Z.				
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells				
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)				
MEDLINE	20499367				
PUBMED	11042152				
REFERENCE	2 (bases 1 to 1218)				
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L., Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J., and Chen,Z.				
TITLE	Human full length cDNA cloned from cd34+ stem cells				
REFERENCE	Unpublished				
AUTHORS	3 (bases 1 to 1218)				
TITLE	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China				
JOURNAL	Location/Qualifiers				
FEATURES	1..1218				
source	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="CBCAPB12"				
	/cell_type="cd34+ stem cells"				
	/tissue_type="blood"				

SOFTWARE: Corlixa Invention Disclosure Database
SEQ ID NO 841
LENGTH: 504
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-43704-841

Query Match 33.8%; Score 402; DB 1; Length 504;
Best Local Similarity 99.6%; Pred. No. 5,2e-197;
Matches 502; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 200 CTTGGGTGTGAGAGCGAATGGCCGATTCACCGGATCCTGAGACCTGATTGAACATCC 259
DB 1 CCCGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGAGACCTGATTGAACATCC 60
QY 260 TCATCCCTGTGTAGACCGGATTCGATATGTGACAGATCTCAAGAAATTTGTATCAACG 319
DB 61 TCATCCCTGTGTAGACCGGATTCGATATGTGACAGATCTCAAGAAATTTGTATCAACG 120
QY 320 TGGCTGAGCAGTGGCTGTGACTCTGACAAATGTAACCTGCAAAATGATGAGTCCCTT 379
DB 121 TGCTGAGCAGTGGCTGTGACTCTGACAAATGTAACCTGCAAAATGATGAGTCCCTT 180
QY 380 ACCTGGCATCATGAGACCTTACAGGCAAGCTACGGTGTGAGAGACCTGATGATGCCG 439
DB 181 ACCTGGCATCATGAGACCTTACAGGCAAGCTACGGTGTGAGAGACCTGATGATGCCG 240
QY 440 TCACCCAGCTAGTCAAAACCATGATGATGATGATGATGATGATGATGATGATGATG 499
DB 241 TCACCCAGCTAGTCAAAACCATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 500 TCTTCCGGAGACGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTG 559
DB 301 TCTTCCGGAGACGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTG 360
QY 560 ACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
DB 361 ACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 620 TGAAGAGTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 679
DB 421 TGAAGAGTCTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 480
QY 680 AGTCTGAGGGGACCGAGAGTCCG 703
DB 481 AGTCTGAGGGGACCGAGAGTCCG 504

RESULT 15
PCT-US01-43704-58
Sequence 58, Application PC/TUS0143704
GENERAL INFORMATION:
APPLICANT: Corlixa Corporation
APPLICANT: Stolck, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madeline Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.561PC
CURRENT APPLICATION NUMBER: PCT/US01/43704
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corlixa Invention Disclosure Database
SEQ ID NO 58
LENGTH: 563
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-43704-58

Query Match 32.9%; Score 391; DB 1; Length 563;

Best Local Similarity 99.8%; Pred. No. 2,5e-191;
Matches 441; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 140 GCGGCGCTCTCTGTGATTTCCCGGAACACCGTGTGATGTTCTGTCGCCACAGAGAG 199
DB 82 GCGGCGCTCTCTGTGATTTCCCGGAACACCGTGTGATGTTCTGTCGCCACAGAGAG 141
QY 200 CTTGGGTGTGAGAGCGAATGGCCGATTCACCGGATCCTGAGACCTGATTGAACATCC 259
DB 142 CCCGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGAGACCTGATTGAACATCC 201
QY 260 TCATCCCTGTGTAGACCGGATTCGATATGTGACAGATCTCAAGAAATTTGTATCAACG 319
DB 202 TCATCCCTGTGTAGACCGGATTCGATATGTGACAGATCTCAAGAAATTTGTATCAACG 261
QY 320 TGGCTGAGCAGTGGCTGTGACTCTGACAAATGTAACCTGCAAAATGATGAGTCCCTT 379
DB 262 TGCTGAGCAGTGGCTGTGACTCTGACAAATGTAACCTGCAAAATGATGAGTCCCTT 321
QY 380 ACCTGGCATCATGAGACCTTACAGGCAAGCTACGGTGTGAGAGACCTGATGATGCCG 439
DB 322 ACCTGGCATCATGAGACCTTACAGGCAAGCTACGGTGTGAGAGACCTGATGATGCCG 381
QY 440 TCACCCAGCTAGTCAAAACCATGATGATGATGATGATGATGATGATGATGATGATG 499
DB 382 TCACCCAGCTAGTCAAAACCATGATGATGATGATGATGATGATGATGATGATGATG 441
QY 500 TCTTCCGGAGACGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTG 559
DB 442 TCTTCCGGAGACGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGCTG 501
QY 560 ACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 581
DB 502 ACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523

Search completed: September 22, 2002, 18:19:49
Job time: 5451 sec

Query Match	48.18;	Score 572;	DB 5;	Length 1309;
Best Local Similarity	99.48;	Pred. NO. 7.2e-285;		
Matches 892;	Conservative	0;	Mismatches 4;	Indels 1;
				Gaps 1

Oy	203	GGGAGTGGAGCCGAATGGGCGGATTCACGGAGATCCGAGGAGCCGTGGTTTGAACATCTCTCA	262
Db	1192	GGGTGTGTGAGCCGAATGGGCGGATTCACGGAGATCCGAGGAGCCGTGGTTTGAACATCTCTCA	1133
Oy	263	TCCCTGTGTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTTCATCAACGTGC	322
Db	1132	TCCCTGTGTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTTCATCAACGTGC	1072
Oy	323	CTGAGCAATGGGCTGTGATCTCTGACAAATTAATCTGCAAAATCGATGGAATCTTTTACC	382
Db	1072	CTGAGCAATGGGCTGTGATCTCTGACAAATTAATCTGCAAAATCGATGGAATCTTTTACC	1013
Oy	383	TGCGCATCATGAGACCTTTCACAAAGCAAGCTACGAGTGTGGAGAGACCCCTGATGTATGCCGTCA	442
Db	1012	TGCGCATCATGAGACCTTTCACAAAGCAAGCTACGAGTGTGGAGAGACCCCTGATGTATGCCGTCA	953
Oy	443	CCCAAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTTNTGGACAAAGTCT	502
Db	952	CCCAAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTTNTGGACAAAGTCT	893
Oy	503	TCCGGGAACGGGAGATCCCTCTCAATGCCAGCATTTGGATGCGCATCAACCAAGCTGCTGACT	562
Db	892	TCCGGGAACGGGAGATCCCTCTCAATGCCAGCATTTGGATGCGCATCAACCAAGCTGCTGACT	833
Oy	563	GCTGGGGATTCGCGTGCCTNCGTTATGATAGATCAAGATATCCATGTG-CCACCCGGGGTG	621
Db	832	GCTGGGGATTCGCGTGCCTNCGTTATGATAGATCAAGATATCCATGTG-CCACCCGGGGTG	773
Oy	622	AAAAGTCTATGCGATGTCAGATGTGAGAGCGACAGCGCGGAAACGGGCCACACTTTAGAG	681
Db	772	AAAAGTCTATGCGATGTCAGATGTGAGAGCGACAGCGCGGAAACGGGCCACACTTTAGAG	713
Oy	682	TCTGAGGAGGACCCAGAGTGTGGCCATCAATGTGGCAGAAAGGAGAAACAGGCCCCAGATC	741
Db	712	TCTGAGGAGGACCCAGAGTGTGGCCATCAATGTGGCAGAAAGGAGAAACAGGCCCCAGATC	653
Oy	742	CTGGGCTTCCGAAGCAGAAAAGGCTGAAACAGATTAATCAGGACAGAGAGAGGCCACGTGCA	801
Db	652	CTGGGCTTCCGAAGCAGAAAAGGCTGAAACAGATTAATCAGGACAGAGAGAGGCCACGTGCA	593
Oy	802	GTTTGTGGCAGAGGCCAGAGGCTTAAAGCTGGAAGCTATTTCGATCCTGGCTGCACCTCTGACA	861
Db	592	GTTTGTGGCAGAGGCCAGAGGCTTAAAGCTGGAAGCTATTTCGATCCTGGCTGCACCTCTGACA	533
Oy	862	CACATATAGAGATGCGAGCAGCTTCACTGACTGTGGCCAGAGCAGTATGTACGCGCGTTTC	921
Db	532	CACATATAGAGATGCGAGCAGCTTCACTGACTGTGGCCAGAGCAGTATGTACGCGCGTTTC	473
Oy	922	TTCGAAACTGGGCAAGAGCTCAACACATCTATCTATCTGCGCTCCACACCTGGCGCATCTCAC	981
Db	472	TTCGAAACTGGGCAAGAGCTCAACACATCTATCTATCTGCGCTCCACACCTGGCGCATCTCAC	413
Oy	982	AGCATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAAGCCCACTGCGCAGG	1041
Db	412	AGCATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAAGCCCACTGCGCAGG	353
Oy	1042	ACTTCAGACTCACTCTCACTGAGTGGGAGCAGCAGAGATGTCCAGGGTATCAGATGCAAGT	1098
Db	352	ACTTCAGACTCACTCTCACTGAGTGGGAGCAGCAGAGATGTCCAGGGTATCAGATGCAAGT	296

RESULT 11
PCT-US01-43704-1161/C
Sequence 1161, Application PC/TUS0143704
GENERAL INFORMATION:
APPLICANT: Corixa Corporation
APPLICANT: Stoik, John A.

```

APPLICANT: Xu, Jiangchun
APPLICANT: Chenaalt, Ruth A.
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.561PC
CURRENT APPLICATION NUMBER: PCT/US01/43704
CURRENT FILING DATE: 2001-11-19
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 1161
LENGTH: 622
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 615
OTHER INFORMATION: n = A,T,C or G
PCT-US01-43704-1161

```

Query Match	42.8%	Score 509	DB 1	Length 622
Best Local Similarity	99.7%	Pred. No.	2.6e-253	
Matches 609	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

OY	321	GCCTGACGAGTCGGCGCTGAGATCTCTGCAAAATGTAACCTGCAAAATCATGAGATCTTTA	380
Db	611	GCCTGACGAGTCGGCGCTGAGATCTCTGCAAAATGTAACCTGCAAAATCATGAGATCTTTA	552
OY	381	CTCTGCGATCATTGACACCTTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGATATGCCGT	440
Db	551	CTCTGCGATCATTGAGACCCCTTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGATATGCCGT	492
OY	441	CACCCAGCTAGCTCAAAACAMCCATGAAGATCAGACCTGGGCAAACTCTCTATGACAAAGT	500
Db	491	CACCCAGCTAGCTCAAAACAMCCATGAAGATCAGACCTGGGCAAACTCTCTGAGCAAAAGT	432
OY	501	CTTCCGGGAAGGGAGATCCCTGATATGCGACATTTGTGGATGGCCATCAACCAAGCTGCTGA	566
Db	431	CTTCCGGGAAGGGAGATCCCTGATATGCGACATTTGTGGATGGCCATCAACCAAGCTGCTGA	372
OY	561	CTGCTGGGGATATCCGCTGCTCCCTGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGGT	620
Db	371	CTGCTGGGGATATCCGCTGCTCCCTGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGGT	312
OY	621	GAAAGAGCTATGTCAGATATGCGAGGTGGAGGCGACGCGGGGAAAGGGGCGACAGTTCTAGA	680
Db	311	GAAAGAGCTATGTCAGATATGCGAGGTGGAGGCGAGGCGGCGGAAAGGGGCGACAGTTCTAGA	232
OY	681	GTCTGAGGGGACCCGAGAGTGGCCATCAATGTGTGCGAAGGGAAGAACAGGCCCGCAT	740
Db	251	GTCTGAGGGGACCCGAGAGTGGCCATCAATGTGTGCGAAGGGAAGAACAGGCCCGCAT	192
OY	741	CCCTGGCCCTCCGAACACAGAAAAAGCTGAAACAGATTAATTCAGGACAGACGAGAGGCGCATGC	800
Db	191	CCCTGGCCCTCCGAACACAGAAAAAGCTGAAACAGATTAATTCAGGACAGAGAGAGGCGCATGC	132
OY	801	AGTTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTGCAATCCTGGCTGCAGCTCTGAC	860
Db	131	AGTTCTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTGCAATCCTGGCTGCAGCTCTGAC	72
OY	861	ACAACATTAATGAGATATCAGACAGCTTACAGCTGATGGCGGACAGATATTCAGCGGTT	920
Db	71	ACAACATTAATGAGATATCAGACAGCTTACAGCTGATGGCGGACAGATATTCAGCGGTT	12
OY	921	CTCCAAACTGG 931	
Db	11	CTCCAAACTGG 1	

RESULT 12
US-10-146-502-193
; Sequence 193, Application US/10146502

PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 106
LENGTH: 1322
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (82)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1320)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1321)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1322)
OTHER INFORMATION: n equals a,t,g, or c
US-10-219-793-106

Query Match 72.0%; Score 855; DB 6; Length 1322;

Best Local Similarity 99.8%; Pred. NO. 0; Mismatches 2; Indels 0; Gaps 0;

DB 139 cggccctcctgctgagatgcccgaacacccgctggtactgtcgtccgacagagagcc 198

202 TGGTGTGAGCAATGGCCGATTCACCGGATCTCTGAGCCTGTGTAACATCTC 261
199 tgggtggtgagcgaatggccgatctccaccgagctctgagcctgttgaaacatcc 258
262 ATCCCTGTAGACCGGATCCGATATGTGACAGTCTCAAGAAATGTATCAACGTC 321
259 atccctgttagaccggatccgatagtcagagcttcaaggaaattgtacataacgctg 318
322 CCTGAGCAGTGGCTGTGACTCTGACAAATGTAATCTGCAATGATGAGTCTTTAC 381
319 cctgagcagctgctgtgactctcgacaaatgtaactctgcaaatcgatgagctcttac 378
382 CTGGGCATCATGACCTTTAAAGGCAAGCTACGCTGTGGAGACCCCTGATGCGCTG 441
379 ctggcatcatgacaccttaacaggcaagctacgctgtggagacccctgagatgacgctc 438
442 ACCCAGCTAGCTCAACCAACCATGAGATCAGAGCTCGCAACACTCTGTTGACAAATGTC 501
439 acccagctagctcaacacacacataagatcagagctcgcaaacctctcttgacaaagtc 498
502 TTCCGGAGACGGAGTCCCTGAATGCCAGATTGTGATGCCATCAACCAAGCTGTGAC 561
499 ttccggagacggagttccctgaatgccagattgtgattgcacataacaaagctgtgac 558
562 TGCTGGGTATCCCTGCTTCGCTTATGATCAAGATATTCATGTCACCCCGGCTG 621
559 tgcgtggatccgcctgctctcgttatagtatcaagatatccatgtagccaccgggctg 618
622 AAAGGTCTATGACATGACAGTGCAGAGCAGAGCGGCAACGGCCACATTTCTAGAG 681
619 aaaggtctatgacatgacagatgcagatgcagagcgagcgaaacggccacagctctagag 678
662 TCTGAGGGGACCCGAGAGTCCGCCATCAATGTGCGAAGAGCAACAAACAGCCCAAGATC 741
679 tctgaggggacccgagagctcgccatcaatgtgagagaaggaagaaacaggcccaagatc 738
742 CTGGCCTCCGAGCAAGGCTGACAGATTAATCAAGGCGAGCGAGAGCCAGTGCA 801
739 ctggcctccgagcagaagctgaaacagataaactcaagcagcgagagagcgcaagtgca 798
802 GTTCTGCGCAAGGCCAAGGCTAAAGCTGAAGCTATTCGAATCTGTGCTGACCTTGACA 861
799 gtctgcgcaagcgcaagcttaagctgaagctatctgaaatctgtgcaagctctgaca 858
862 CAACATATGAGATGACACACTTCACTGACTGTGGCGGACAGATATGACGCCGTTTC 921
859 caacataatgagatgcagcaagcttcaactgactgtggtcgagacagatgtcagcgagctc 918
922 TCCAAACTGGCCAGGAGCTCAACACTATCTACTGACCCCTCCACCCCTGGCATGTACAC 981
919 tccaaactggccagagacatcccaactatccctacagctcccaacccttgagatgtcacc 978
982 AGCATGTGCTCAGGCGCATGGGTGTATGTGAGAGCCCTCACCACCAAGCCCACTGCCAGGG 1041
979 agcatgtgtgctcagcgcatgtgtgtatatatgagccctcacaagccccagtgcaagg 1038
1042 ACTCCAGACTACTCTCCAGTGGAGACACAGAGATGTCAGAGGTCAAGGTCAATGCAAGT 1098
1039 actccagactactctccagtggtgagcagagagatgtccaggtgtacagatgcaagt 1095

RESULT 7

US-09-629-469A-11009
Sequence 11009, Application US/09629469A

GENERAL INFORMATION:

APPLICANT: OTA, TOSHIO

APPLICANT: ISOGAI, TAKAO

APPLICANT: NISHIKAWA, TETSUO

APPLICANT: HAYASHI, KOJI

APPLICANT: SAITO, KAORI

APPLICANT: YAMAMOTO, JUNICHI

APPLICANT: ISHII, SHIZUKO

Db 920 tctccaaactgagccaaaggaactccaacactctcactgcccctccaacccctgagcatgca 979
Qy 980 CCAGCATGTGTGGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAGAAAGCCCACTGGCAG 1039
Db 980 ccagcatggtgagctcagagccatggtgtatataatgagagccctccaacaaagcccccagtgccag 1039
Qy 1040 GGACTCCAGACTCAGCTTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCANGT 1098
Db 1040 ggaactccagactcactctccagctgagagcagagagatgctcaggggtacagatgcaagt 1098

RESULT 5

US-10-119-428-37
Sequence 37, Application US/10119428
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Wehrman, Tom
APPLICANT: Ren, Felyan
APPLICANT: Ma, Yungling
APPLICANT: Zhou, Ping
APPLICANT: Zhao, Qiang A.
APPLICANT: Yang, Yonghong
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 789CIP2
CURRENT APPLICATION NUMBER: US/10/119,428
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 37
LENGTH: 9098
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (75)..(9098)
US-10-119-428-37

Query Match 72.18; Score 857; DB 7; Length 9098;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCGCGCGCTCTCTGATGATCCCGGAACACCGGTACTGTCGTCGCCGACGAGGAGG 199
Db 151 gccgcgcctctctgcatctgccccgaacacgcgtgtaactctcgctgcgcagagaagg 210
Qy 200 CCTGGGTGTGAGCAGATGGCCGATTCACCGGATCTGAGCCCTGGTTGAACATCC 259
Db 211 cctgggtgtgagcagatggccgatctcacccgatcccgagcccggttgtaactcc 270
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGACAGACTCTCAAGAAATTGTATCAACG 319
Db 271 tcatccctgtgtagaccggaatccgatatgtgcagagttccaagaaattgtcatcaag 330
Qy 320 TGGCTAGAGAGTGGGCTGTACTCTGACAAATGTAATCTGCAAAATGATGGAGTCTT 379
Db 331 tggctagagagtggtgtgactctcgacaaatgtaactctgcaaatcgatgaggtccctt 390
Qy 380 ACCTGGCATCATGAGCCCTTACAGCAAGCTACGGTGTGAGGAGCCCTGAGATGCGG 439
Db 391 acctgcatcatgagcccttacaagcaagctacggtgtgagagaccctgagatagcg 450

Qy 440 TCACCCAGCTAGCTCAAAACACATGAGATCAGAGCTCGGCAAACTCTNTGACAAAG 499
Db 451 tcaaccagctagctcaaaacacatgagatcagagcttcgcaaacctctcttgaaag 510
Qy 500 TCTTCCGGGAACGGGACTCCCTGAAATGCCAGCATTTGTGATGGCATTAACCAAGCTCTG 559
Db 511 tcttcgggaacgggagctccctgaaatgccagcatltgtgagatgcataacaaagctgtcg 570
Qy 560 ACTGCTGGGTATCCGCTGCTGCTTTCATGATGATCAAGGATATTCATGATGATGATGATG 619
Db 571 actgctgggtatccgctgctgctcctcgtatgagatcaagatatcatatgtgtccaccggg 630
Qy 620 TGAAGAGCTTATGACATCAGGTGAGGACAGCGCGGAAAGCGGCCACAGTTCTAG 679
Db 631 tgaagagcttatgacatgcaagtgagagcagagcgcggaacagggccacagttctag 690
Qy 680 AGCTGAGGGAGCCGAGCTCGGCTCATGANTGGCAGAGAAAGGAACAAACAGGCCAGA 739
Db 691 agctgagggagcccgagctcggtccatcaatgtgagagaagggaagaaacagccaga 750
Qy 740 TCTGAGCTCCGAAGCAGAAAGGCTGACAGATTAATACAGCAGCAGAGAGCCAGTG 799
Db 751 tctgagctccgaagcagaaagctgtacaagataaatacagcagagagagccagtg 810
Qy 800 CAGTTCTGGCGAAGCCAGGCTTAAGCTAAGCTATTCGAATCTGCTGACAGCTCTGA 859
Db 811 cagttctggcgaaagccagagcttaagcttaagctatctgcatctgtgcagctctga 870
Qy 860 CACAACATTAATGGAGTGCAGAGCTTCACTGAGTGGCGAGCATATGTACGGCGCT 919
Db 871 cacaacataatgagatgagcagcttcacagactgagcagagctatagcaagcgcgt 930
Qy 920 TCTCCAACTGGCCAAAGAGTCCACACTATCTACTGCCCCCAACCTCGGGAGTGA 979
Db 931 tctccaaactggccaagagactccaactatctctactgcctcccaaccctggagatgca 990
Qy 980 CCAGCATGTGTGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAGAAAGCCCACTGGCAG 1039
Db 991 ccagcatggtgagctcagagccatggtgtatataatgagagccctccaacaaagcccccagtgccag 1050
Qy 1040 GGACTCCAGACTCAGCTTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCANGT 1098
Db 1051 ggaactccagactcactctccagctgagagcagagagatgctcaggggtacagatgcaagt 1109

RESULT 6

US-10-219-793-106
Sequence 106, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
PRIOR FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608

Db 679 cctgggtgtagcgaatgggcccgtccacccggtccctggagccttggtgaacatcc 738
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGTCATCAACG 319
Db 739 tcattccctgtgttaacccggatccgatattgtgcagaagctccaagaaattgtcacaacg 798
Qy 320 TGCCTGACAGCTGGCTGTGACTCTGCACAAATGTACTCTGCAAAATCGATGGAGTCTTT 379
Db 799 tgcctgagcagtcggtcgtgactctgcacaatctgcaaatcgtatgagtcgctt 858
Qy 380 ACCCTGGATCATGTGACCTTTCAGAGCAAGTACGGTGTGAGAGACCTGTGATGCCG 439
Db 859 accctgcacatctgagcccttacaagcaagctacggtgtgagagccctgtcgtatgcg 918
Qy 440 TCACCAGCTACCTCAAAACCATGAGATCAGAGCTGGCAACTCTCTNNGCACAAAG 499
Db 919 tcaaccagctgctcaacaacacatgagatcagagctgcgcaaatctctcctgcgacaag 978
Qy 500 TCTTCGGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGTG 559
Db 979 tcttcgggaacgggagtcctcgtgaatgcccagatctgtgagtcacatcaaccagctgtg 1038
Qy 560 ACTGCTGGGATTCGGCTGCTTNCCTTATGATCAAGATTCATGTGCCACCCGGG 619
Db 1039 actgctgggtatccgctgctccgttatgagatcaagataccatgtgcccacccggg 1098
Qy 620 TGAAGAGTCTATGAGATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCTT 679
Db 1099 tgaagagctctcagatgagagtcggaagcagagcgaggaacgggacacagcttctg 1158
Qy 680 ACTTGAGGAGCCCGAGAGTGGCCATCATATGTGGCAGAAAGGAAACAGGCCACAGA 739
Db 1159 agctcgaaggagcccgagagctgcgcacatcatgtgagcagaaggagaaacggccaga 1218
Qy 740 TCCCTGGCTCCCAACAGAAAGCTGTAACATTAATAGGACAGAGAGAGGACGATG 799
Db 1219 tccctggcctccgaagcagaagagctgaaacagataaatacgaagagagagcgccagtg 1278
Qy 800 CAGTTCTGGGAGGAGCAAGGCTAAAGCTGAAGCTATTGCAATCCGCTGAGCTCTGA 859
Db 1279 cagttctggcgaaggcagaagctaaagctgaagctatctgaaatccctgctgcagcttga 1338
Qy 860 CACAACATATGAGATGACAGAGCTTCACTGACTGTGGCCGAGCATATGTCAGCGCT 919
Db 1339 cacaacataatgagatgagcagctcactgactgtgagcagagatgtcagcgct 1398
Qy 920 TCTCCAACTGGCCAAAGACTCCAACTATCTACTGCGCTCCAAAGCCCTGGCGATGCA 979
Db 1399 tctccaaactggcgaagactccaacatacctcactgctccaaacccctggagatgta 1458
Qy 980 CCAGATGGTGGCTCAGGCAATGGGTATATGAGAGCCCTCAACCAAGCCCAAGTCCAG 1039
Db 1459 ccagatggctgctcagcagcagtgatataatgagcctccaccaaagcccaagtcag 1518
Qy 1040 GCAGCTCAGACTCTCTCCAGTGGGAGCAGCAGAGATGTCAGGGTACAGATGCAAGT 1098
Db 1519 ggaactcagactcactcctcagtgagagcagagagatgtccagggtaacagatgcaagt 1577

RESULT 4
US-10-198-846-12975
; Sequence 12975, Application US/10198846
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION OF BREAST CANCER ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12975
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1, 2, 1909, 1910
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12975

Query Match 72.1%; Score 857; DB 6; Length 1910;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGCGCTCTCTGTGATGCCCCGAAACACCGTGTACTGTTCTGCCGACAGAGAG 199
Db 140 gccgcgctctctgtgatgccccgaaacacggtgtactgttctgcccagcagag 199
Qy 200 CCTGGGTGATGAGCGAATGGCGCATTCACCGGATCTGAGACCTGGTTGAACATCC 259
Db 200 cctgggtgatgagcgaatggcgcatctcacccggtatcctggagccttgatgaacatcc 259
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGTCATCAACG 319
Db 260 tcattccctgtgttagaccggatccgatattgtgcagaagctccaagaaattgtcacaacg 319
Qy 320 TGCCTGACAGCTGGCTGTGACTCTGCACAAATGTACTCTGCAAAATCGATGGAGTCTTT 379
Db 320 tgcctgacagctggcgtgtgactctgcacaatctgcaaatcgtatgagtcgctt 379
Qy 380 ACTGCTGGGATTCGGCTGCTTNCCTTATGATCAAGATTCATGTGCCACCCGGG 439
Db 380 actgctgggtatccgctgctccgttatgagatcaagataccatgtgcccacccggg 439
Qy 440 TCACCAGCTACCTCAAAACCATGAGATCAGAGCTGGCAACTCTCTNNGCACAAAG 499
Db 440 tcaaccagctgctcaacaacacatgagatcagagctgcgcaaatctctcctgcgacaag 499
Qy 500 TCTTCGGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGTG 559
Db 500 tcttcgggaacgggagtcctcgtgaatgcccagatctgtgagtcacatcaaccagctgtg 559
Qy 560 ACTGCTGGGATTCGGCTGCTTNCCTTATGATCAAGATTCATGTGCCACCCGGG 619
Db 560 actgctgggtatccgctgctccgttatgagatcaagatacctcctgcagccccggg 619
Qy 620 TGAAGAGTCTATGAGATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCTT 679
Db 620 tgaagagctctcagatgagagtcggaagcagagcgagagcgaggaacgggacacagtcag 679
Qy 680 AGTCTGAGGAGCCGAGAGTGGCCATTAATGTGCGACAGAGGAAAGAAACAGGCCACA 739
Db 680 agtctgagggagcccgagagctcgccatcatatgtgacgaaggagaaacagggccaga 739
Qy 740 TCTTCGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 799
Db 740 tcttcggctccgagagcagaagagctggaacagataacagagcagagagagccagtg 799
Qy 800 CAGTTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 859
Db 800 cagttctggcgaaggcagaagcctaagcctaagctcagatccctgcgtcagcgtcgtga 859
Qy 860 CACAACATATGAGATGACAGAGCTTCACTGACTGTGGCCGAGCATATGTCAGCGCT 919
Db 860 cacaacataatgagatgagcagctcactgactgtgagcagagatgtatgtaagcgct 919
Qy 920 TCTCCAACTGGCCAAAGACTCCAACTATCTACTGCGCTCCAAAGCCCTGGCGATGCA 979
Db 920 tctccaaactggcgaagactccaacatacctcactgctccaaacccctggagatgta 979


```

: PRIOR APPLICATION NUMBER: 60/060,844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 45
: LENGTH: 1337
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1335)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1336)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1337)
: OTHER INFORMATION: n equals a,t,g, or c
US-10-219-793-45
```

Query Match 72.1%: Score 857; DB 6; Length 1337;

Best Local Similarity 99.8%: Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 140 GCCGCCCTCTCTGATTGCCCGAAGACCGTGTACTGTTGCGCCAGCAGAGG 199
    |||||||
Db 152 gccgcgctctctctgtgattgccccgaacacgctgttactgttcgtccgacagagag 211
Oy 200 CCTGGTGTGAGCGAATGGCCGATTCCACCGATCCTGAGACCTGTTGAACATCC 239
    |||||||
Db 212 ccctgggtgtgagcgaaatggcgatctccacggatctcctggagcctgttgaaatcc 271
Oy 260 TCATCCCTGTGTAGACCGGATTCGATGTGTGACAGTCTCAAGAAATTTGTATCAAG 319
    |||||||
Db 272 tcatccctgtgttagacggatccgatactgtgagagctccaaggaattgtatcaacg 331
Oy 320 TCCTGAGAGTGTGAGTGTGACTCTGACAAATGTAATGATGAGTGCCTT 379
    |||||||
Db 332 tgcctgagagcagctgtgacctctgacaatgttaactctgcaatcgatgggtcctt 391
Oy 380 ACCTGGCATCATGAGACCTTACAGCAAGCTAGCGTGTGAGGACCTGAGTATGCCG 439
    |||||||
Db 392 acctgcatcatgagaccttaacaagcaagctagctgtggagacccctgagctatgccc 451
Oy 440 TCACCCAGCTAGCTCAAAACCATGATGATGATGATGATGATGATGATGATGATGATG 499
    |||||||
Db 452 tcacccaagctagctcaacaacacagatcagatcagatcagatcagatcagatcagatc 511
Oy 500 TCTTCCGAGACGAGTCCCTGAATGACGAGTGTGATGATGATGATGATGATGATGATG 559
    |||||||
Db 512 tcttcgaggaaagggagctccctgaaatgtgcaatgtgcaatgtgcaatgtgcaatgtgca 571
Oy 560 ACTGCTGGGATTCGCTGCTTCGCTTATGATGATGATGATGATGATGATGATGATGATG 619
    |||||||
Db 572 actgctgggggtatcgcgtctcctcgtctatgagatcaaggtatcaatgtgcaatgtgca 631
Oy 620 TGAAGAGCTATGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 679
    |||||||
Db 632 tgaagaagctctatgcaatgcaaggtgagagcagagcagagcagagcagagcagagcag 691
Oy 680 ACTTGTAGAGGAGACCGGAGATCGGATCATGATGATGATGATGATGATGATGATGATG 739
    |||||||
Db 692 agtctgagagggagcccgagagctgggccaatgtgagcaagaaggaagaaacagagccca 751
```

```

Oy 740 TCCTGCGCTCCGAGCAGAAAGCTGMAACGATTAATCAGGACGAGAGGCCAGTG 799
    |||||||
Db 752 tcttgccctcgaagcagaaagctgaaacagataatcagacagcagagagagccagctg 811
Oy 800 CAGTTTGGCGAAGGCCAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTA 859
    |||||||
Db 812 cagttctggagagggcgaagcctaagctgaagctatcgaatctcctgctgctgctgca 871
Oy 860 CACAACATAATGAGATGCGAGCTTCACTGACTGTGGCGGAGAGATGATGAGCGCT 919
    |||||||
Db 872 cacaacataatgagatgacgacgcttcaactgactgtgctgagcagatgacgagctg 931
Oy 920 TCTCCAAACTGGCCCAAGACTCCACACATATCTACTGCTCCCTCAACCTGCGGATGCA 979
    |||||||
Db 932 tctccaactgccaagactccaacactatctactgctccctccaacacctgtgcagtca 991
Oy 980 CCAGCATGTGTGCTCAGGCGCATGGGTATATATGAGAGCCCTCACCAGAGCCAGTGC 1039
    |||||||
Db 992 ccagcatgtgtctcagcgcaatgtgtatagagccctcaccacaaagcccgctcag 1051
Oy 1040 GGAATCCAGACTCACTCTCCAGTGGGAGCAGAGATGTCAGGATGATGATGATGATG 1098
    |||||||
Db 1052 ggaatccagactcaactctcctcagctgagagcagagatgtccaggttaccagatgca 1110
```

RESULT 3

US-09-785-276A-24973

; Sequence 24973, Application US/09785276A

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Endege, Wilson

; APPLICANT: Monahan, John

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF

; TITLE OF INVENTION: HUMAN PROSTATE CANCER

; FILE REFERENCE: MRI-007B

; CURRENT APPLICATION NUMBER: US/09/785,276A

; PRIORITY FILING DATE: 2001-02-16

; PRIORITY FILING DATE: 2000-02-17

; PRIORITY FILING DATE: 2000-02-17

; PRIORITY FILING DATE: 2000-03-16

; PRIORITY FILING DATE: 2000-03-16

; PRIORITY FILING DATE: 2000-05-25

; PRIORITY FILING DATE: 2000-05-25

; PRIORITY FILING DATE: 2000-06-09

; PRIORITY FILING DATE: 2000-06-09

; PRIORITY FILING DATE: 2000-07-18

; PRIORITY FILING DATE: 2000-07-18

; PRIORITY FILING DATE: 2000-12-13

; PRIORITY FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 62232

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 24973

; LENGTH: 1842

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1841..1842

; OTHER INFORMATION: n = A,T,C or G

US-09-785-276A-24973

Query Match 72.1%: Score 857; DB 5; Length 1842;

Best Local Similarity 99.8%: Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 140 GCCGCCCTCTCTGATTGCCCGAAGACCGTGTACTGTTGCGCCAGCAGAGG 199
    |||||||
Db 619 gccgcgctctctctgtgattgccccgaacacgctgttactgttcgtccgacagagag 678
Oy 200 CCTGGTGTGAGCGAATGGCCGATTCCACCGATCTGAGAGCTGTTGAACATCC 259
    |||||||
```


[illegible]

```

Db      1021  |||||||
          ACCAAGGCCAGGCCAGGAGCTCCAGCTCACTCCAGTGGGAGCAGCAGATATCTC 1080
QY      1081  CAGGCTACAGATGCAGAGTNTTGGATGAGAGCACTTGATCGAGTCAAGATGAGTACTGAGAC 1140
Db      1081  CAGGCTACAGATGCAGAGTNTTGGATGAGAGCACTTGATCGAGTCAAGATGAGTACTGAGAC 1140
QY      1141  TGGGCTTGGCCAGGAGGAGTCTGGGGACACAGACACATTTTCTGATT 1188
          |||||||
          TGGGCTTGGCCAGGAGGAGTCTGGGGACACAGACACATTTTCTGATT 1188
Db      1141  TGGGCTTGGCCAGGAGGAGTCTGGGGACACAGACACATTTTCTGATT 1188

RESULT      2
US-10-219-793-45
; Sequence 45, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1C1
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 16:48:58 : Search time 322.14 Seconds
(without alignments)
11832.429 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCTTCTGGAGACNACCGCT.....GGAAGACAGATTTCCTGATT 1188

Scoring table:
OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2138461 seqs, 1604250230 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4276922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/US06_NEM_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEM_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEM_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEM_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEM_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	1188	US-09-898-216-2	Sequence 2, Appl
2	857	72.1	1337	US-10-219-793-45	Sequence 45, Appl
3	857	72.1	1842	US-09-785-276A-24973	Sequence 24973, A
4	857	72.1	1910	US-10-198-846-12975	Sequence 12975, A
5	857	72.1	9098	US-10-119-428-37	Sequence 37, Appl
6	855	72.0	1322	US-10-219-793-106	Sequence 106, Appl
7	806	67.8	1244	US-09-629-469A-11009	Sequence 11009, A
8	674	56.7	1234	US-09-703-256A-3468	Sequence 3468, Ap
9	616	51.9	1150	US-10-198-846-13409	Sequence 13409, A
10	572	48.1	1309	US-09-919-002-11863	Sequence 11863, A
11	509	42.8	622	PCT-US01-43704-1161	Sequence 1161, Ap
12	454	38.2	454	US-10-146-502-193	Sequence 193, App
13	453	38.1	550	US-10-214-403-2895	Sequence 2895, Ap
14	402	33.8	504	PCT-US01-43704-841	Sequence 841, App
15	391	32.9	563	PCT-US01-43704-58	Sequence 58, Appl
16	390	32.8	448	US-09-920-300A-1075	Sequence 1075, Ap
17	390	32.8	448	US-10-099-926-1075	Sequence 1075, Ap
18	357	30.1	357	US-10-214-403-2873	Sequence 2873, Ap
19	352	29.6	439	PCT-US02-25766-1798	Sequence 1798, Ap
20	331	27.9	691	US-10-097-105-1428	Sequence 1428, Ap
21	316	26.6	445	US-09-918-995-5335	Sequence 5335, Ap
22	279	23.5	362	US-09-721-544-8507	Sequence 8507, Ap
23	267	22.5	478	US-09-918-995-31316	Sequence 31316, A
24	262	22.1	430	US-09-785-276A-37723	Sequence 37723, A
25	253	21.5	235	US-09-539-331D-4825	Sequence 4825, Ap

C	26	250	21.0	301	1	PCT-US02-12378-26	Sequence 26, Appl
C	27	250	21.0	301	7	US-10-124-805-26	Sequence 26, Appl
	28	221	18.6	221	5	US-09-539-331D-15714	Sequence 15714, A
	29	221	18.6	300	5	US-09-539-331D-21791	Sequence 21791, A
	30	216	18.2	253	5	US-09-539-331D-688	Sequence 688, App
	31	211	17.8	211	5	US-09-540-210B-18465	Sequence 18465, A
	32	199	16.8	199	5	US-09-540-210B-6996	Sequence 6996, Ap
	33	171	14.4	567	5	US-09-629-469A-1950	Sequence 1950, Ap
	34	169	14.2	740	7	US-10-097-105-596	Sequence 596, App
	35	160	13.5	239	5	US-09-539-331D-16278	Sequence 16278, A
	36	151	12.7	232	5	US-09-539-331D-12906	Sequence 12906, A
	37	149	12.5	300	5	US-09-539-331D-9558	Sequence 9558, A
	38	148	12.5	227	5	US-09-539-331D-15486	Sequence 15486, A
	39	147	12.4	282	5	US-09-540-210B-15126	Sequence 15126, A
	40	144	12.1	214	5	US-09-540-210B-7087	Sequence 7087, Ap
	41	139	11.7	265	5	US-09-539-331D-14079	Sequence 14079, A
	42	125	10.5	390	5	US-09-785-276A-7770	Sequence 7770, Ap
	43	124	10.4	902	6	US-10-198-846-6734	Sequence 6734, Ap
	44	122	10.3	242	5	US-09-539-331D-30091	Sequence 30091, A
C	45	122	10.3	1158	6	US-10-198-846-5820	Sequence 5820, Ap

ALIGNMENTS

RESULT 1
US-09-898-216-2
Sequence 2, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898, 216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1188 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-898-216-2

```

: PRIOR FILLING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: US 09/248,797
: PRIOR FILLING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: US 09/271,490
: PRIOR FILLING DATE: 1999-03-18
: PRIOR APPLICATION NUMBER: US 09/293,972
: PRIOR FILLING DATE: 1999-04-15
: PRIOR APPLICATION NUMBER: US 09/274,861
: PRIOR FILLING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: US 60/125,453
: PRIOR FILLING DATE: 1999-03-19
: PRIOR APPLICATION NUMBER: US 60/126,605
: PRIOR FILLING DATE: 1999-03-26
: PRIOR APPLICATION NUMBER: US 09/306,350
: PRIOR FILLING DATE: 1999-05-07
: PRIOR APPLICATION NUMBER: US 09/399,720
: PRIOR FILLING DATE: 1999-09-21
: PRIOR APPLICATION NUMBER: US 09/404,284
: PRIOR FILLING DATE: 1999-09-21
: PRIOR APPLICATION NUMBER: US 09/465,877
: PRIOR FILLING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
: PRIOR FILLING DATE: 1999-01-19
: NUMBER OF SEQ ID NOS: 10410
: SOFTWARE: PC_CT_genes Version 1.02
: SEQ ID NO: 8410
: LENGTH: 1398
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1296)...(264)
: OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
US-09-496-914A-8410

```

```

Query Match          72.1%; Score 857; DB 18; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 140 GCGGCGCTCCCTGCGATGCGCCGCAACACCGTGCTCTGCTGCGCGGACGAGG 199
Db 1254 GCGGCGCTCCCTGCGATGCGCCGCAACACCGTGCTCTGCTGCGCGGACGAGG 1195

Oy 200 CCGGCGGTGCGAGCGAATGCGCGCATTCACCGGATCCTGAGCGCTGTTGAACATCC 259
Db 1194 CCGGCGGTGCGAGCGAATGCGCGCATTCACCGGATCCTGAGCGCTGTTGAACATCC 1135

Oy 260 TCATCCCTGTGTAGACCGGATCGATATGTGACAGAGTCTCAAGGAATGTTCATCAACG 319
Db 1134 TCATCCCTGTGTAGACCGGATCGATATGTGACAGAGTCTCAAGGAATGTTCATCAACG 1075

Oy 320 TGCCGAGACAGCGCGCTGCGACTGCGACAATGTAACCTGCAATGATGAGTCCTT 379
Db 1074 TGCCGAGACAGCGCGCTGCGACTGCGACAATGTAACCTGCAATGATGAGTCCTT 1015

Oy 380 ACCTCGCGATCATGAGACCTTACAGGCAAGCTAGCGTGTGAGGACCCCTGATGCGG 439
Db 1014 ACCTCGCGATCATGAGACCTTACAGGCAAGCTAGCGTGTGAGGACCCCTGATGCGG 955

Oy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGCGCAACTCTCTNTGACAAG 499
Db 954 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGCGCAACTCTCTNTGACAAG 895

Oy 500 TCTTCGCGAAGCGGAGTCCCTGGAATGCGAGATGCGATGCGATGCGATGCGATGCG 559
Db 894 TCTTCGCGAAGCGGAGTCCCTGGAATGCGAGATGCGATGCGATGCGATGCGATGCG 835

Oy 560 ACTGCTGGGGTATCGCGTCCCTGCTATGAGATCAAGGATATCCATGTCACCCCGG 619
Db 834 ACTGCTGGGGTATCGCGTCCCTGCTATGAGATCAAGGATATCCATGTCACCCCGG 775

```

```

Oy 620 TGAAGACTCTATGCGATGCGAGTGAGGAGAGCGCGGCAAAAGGCGCACAGTTCTAG 679
Db 774 TGAAGACTCTATGCGATGCGAGTGAGGAGAGCGCGGCAAAAGGCGCACAGTTCTAG 715

Oy 680 AGCTGAGGGGACCCGAGAGTTCGGCCATCATGTGTGCGAAGGGAAGAACAGGCCCA 739
Db 714 AGCTGAGGGGACCCGAGAGTTCGGCCATCATGTGTGCGAAGGGAAGAACAGGCCCA 655

Oy 740 TCTCGGCTCCGAGCAGAAAAAGGCTGAACAGATTAATTCAGGACGACAGAGGCGAGTG 799
Db 654 TCTCGGCTCCGAGCAGAAAAAGGCTGAACAGATTAATTCAGGACGACAGAGGCGAGTG 595

Oy 800 CAGTTCTGCGCAAGGCGCAAGGCTAAAGCTGAAGCTATTGCAATCCGTGCGACGCTGA 859
Db 594 CAGTTCTGCGCAAGGCGCAAGGCTAAAGCTGAAGCTATTGCAATCCGTGCGACGCTGA 535

Oy 860 CACAACATTAATGAGATCAGACAGCTTACACGACTGTGTGCGGAGAGATATGTCAAGCGGT 919
Db 534 CACAACATTAATGAGATCAGACAGCTTACACGACTGTGTGCGGAGAGATATGTCAAGCGGT 475

Oy 920 TCTCCAACTGCGCAAGGACTGCAACACTATCTACTGCTCCCAACCTGCGGATGCA 979
Db 474 TCTCCAACTGCGCAAGGACTGCAACACTATCTACTGCTCCCAACCTGCGGATGCA 415

Oy 980 CCAGCATGCTGCTCAGGCGCATGGGTGTATATGAGCCCTCACCAGGCGGAGTCCAG 1039
Db 414 CCAGCATGCTGCTCAGGCGCATGGGTGTATATGAGCCCTCACCAGGCGGAGTCCAG 355

Oy 1040 GAGCTCCGAGTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAT 1098
Db 354 GAGCTCCGAGTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

```

Search completed: September 22, 2002, 18:14:09
Job time: 6820 sec

```

; EARLIER APPLICATION NUMBER: US 09/234,611
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 09/240,371
; EARLIER FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 09/277,227
; EARLIER FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: US 09/271,490
; EARLIER FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 09/293,972
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/274,861
; EARLIER FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: US 60/125,453
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: US 60/126,605
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: US 09/306,350
; EARLIER FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/399,720
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 09/404,284
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER:
; EARLIER FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 10451
; SOFTWARE: PC_CT_genes Version 1.0
; SEQ ID NO: 865
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1296)...(264)
; OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-471-275-865
```

```

Query Match          72.1%  Score 857;  DB 18;  Length 1398;
Best Local Similarity 99.8%  Pred. No. 0;
Matches 957;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;
```

```

QY 140 GCCGGCCCTCTCTGGATTGCCCGGAACACCGTGCTGTTCTGCGCAGCAGAGG 199
    |||||
DB 1254 GCCGGCCCTCTCTGGATTGCCCGGAACACCGTGCTGTTCTGCGCAGCAGAGG 1195
    |||||
QY 200 CTTGGGTGGTGGAGGAATGGGCCGATTCACCGGATCCTGAGCCTGGTTGAACATCC 259
    |||||
DB 1194 CTTGGGTGGTGGAGGAATGGGCCGATTCACCGGATCCTGAGCCTGGTTGAACATCC 1135
    |||||
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGAAATTGTCATCAACG 319
    |||||
DB 1134 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGAAATTGTCATCAACG 1075
    |||||
QY 320 TGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGATGAGTCTTT 379
    |||||
DB 1074 TGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGATGAGTCTTT 1015
    |||||
QY 380 ACCTGCGATCATGAGACCTTACAAAGGCAAGCTAGCGTGTGAGAGACCTTGATGCCG 439
    |||||
DB 1014 ACCTGCGATCATGAGACCTTACAAAGGCAAGCTAGCGTGTGAGAGACCTTGATGCCG 955
    |||||
QY 440 TCACCGACTACCTCAAAACCAATGAGATGATGAGCTGGGCAAACTCTCTNNAGCAAG 499
    |||||
DB 954 TCACCGACTACCTCAAAACCAATGAGATGATGAGCTGGGCAAACTCTCTNNAGCAAG 895
    |||||
QY 500 TCTTCCGGGAAGGAGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 559
    |||||
DB 894 TCTTCCGGGAAGGAGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 835
    |||||
QY 560 ACTGCTGGGGTATCCGCTCCCTNCGTTATGAGATCAAGGATATTCATGTGCCACCCGGG 619
    |||||
DB 834 ACTGCTGGGGTATCCGCTCCCTNCGTTATGAGATCAAGGATATTCATGTGCCACCCGGG 775
    |||||
```

```

QY 620 TGAAGAGTCTATGACATGACAGTGGAGGAGGAGGCGGAACGGGCCACAGTTCTAG 679
    |||||
DB 774 TGAAGAGTCTATGACATGACAGTGGAGGAGGAGGCGGAACGGGCCACAGTTCTAG 715
    |||||
QY 680 AGTGTGAGGGGACCCGAGAGTGGGCCATCATGTGTGGCAGAGGGAAGAAACAGGCCACA 739
    |||||
DB 714 AGTGTGAGGGGACCCGAGAGTGGGCCATCATGTGTGGCAGAGGGAAGAAACAGGCCACA 655
    |||||
QY 740 TCTTGCCCTCCCAAGCAGAAAGGCTGAACACATTAATAGCAGCAGAGAGAGGCCACAG 799
    |||||
DB 654 TCTTGCCCTCCCAAGCAGAAAGGCTGAACACATTAATAGCAGCAGAGAGAGGCCACAG 595
    |||||
QY 800 CAGTTCTGGCGAAGGCCAAGCTAAAGCTGAAGCTATTGCAATCCGCTGCAGCTCTGA 859
    |||||
DB 594 CAGTTCTGGCGAAGGCCAAGCTAAAGCTGAAGCTATTGCAATCCGCTGCAGCTCTGA 535
    |||||
QY 860 CACAACATATAGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGATATGTCAGCGCT 919
    |||||
DB 534 CACAACATATAGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGATATGTCAGCGCT 475
    |||||
QY 920 TCTCCAACTGGCCCAAGGACTCCACACATATCTACTGCTCCAACTGGCCGATGTGA 979
    |||||
DB 474 TCTCCAACTGGCCCAAGGACTCCACACATATCTACTGCTCCAACTGGCCGATGTGA 415
    |||||
QY 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAAGGCCAGTGCAG 1039
    |||||
DB 414 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAAGGCCAGTGCAG 355
    |||||
QY 1040 GGACTTCACAGCTACTCTCACTGAGGAGCAGCAGCAGATGTCCAGGATACAGATGCAAGT 1098
    |||||
DB 354 GGACTTCACAGCTACTCTCACTGAGGAGCAGCAGCAGATGTCCAGGATACAGATGCAAGT 296
    |||||
```

RESULT 15

```

US-09-496-914A-8410/C
; Sequence 8410, Application US/09496914A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Conligns Obtained
; FILE REFERENCE: From Various Libraries
; FILE REFERENCE: 787
; CURRENT APPLICATION NUMBER: US/09/496,914A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 09/045,400
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: US 09/131,598
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 09/431,517
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: US 09/328,351
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 09/332,782
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/235,076
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: US 09/234,611
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 09/346,956
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: US 09/362,510
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 09/240,371
```

QY 680 AGCTGAGGGACCCGAGAGTCGCCATCATGTGGCGAAGGGAAGAAACAGGCCCCAGA 739
DB 714 AGCTGAGGGAGCCCGAGAGTCGCCATCATGTGGCGAAGGGAAGAAACAGGCCCCAGA 655
QY 740 TCCTGGCCCTCGAAGCACAAGGCTGAACAGATTAATCAGGACACAGAGAGGCCACTG 799
DB 654 TCCTGGCCCTCGAAGCACAAGGCTGAACAGATTAATCAGGACACAGAGAGGCCACTG 595
QY 800 CAGTTCTGGCGAAGCCGAAGCTTAAGCTGAAGCTATTTCGATCCCTGGCTGACGCTGTA 859
DB 594 CAGTTCTGGCGAAGCCGAAGCTTAAGCTGAAGCTATTTCGATCCCTGGCTGACGCTGTA 535
QY 860 CACAACATATGAGATGAGATGAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCACGGCT 919
DB 534 CACAACATATGAGATGAGATGAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCACGGCT 475
QY 920 TCCTCAAACTGGCCGAAGAGCTCCACACTATCTGCTGCTCCCAACCTGGCGATGTCA 979
DB 474 TCCTCAAACTGGCCGAAGAGCTCCACACTATCTGCTGCTCCCAACCTGGCGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGCCAGTGCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGCCAGTGCAG 355
QY 1040 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 354 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 13
PCT-US01-04098A-2212/c
; Sequence 2212, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 2212
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-04098A-2212

Query Match 72.1%; Score 857; DB 1; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCCTCTCTGATTGCCCGAAGACCGTGTACTGTCTGCGCAGCAGAGAG 199
DB 1254 GCCGCGCCTCTCTGATTGCCCGAAGACCGTGTACTGTCTGCGCAGCAGAGAG 1195

QY 200 CCTGGGTGGTGGAGGAAATGGGCGCGATTCCACCGGATTCCTGGAGCTGGTTGAACATCC 259
DB 1194 CCTGGGTGGTGGAGGAAATGGGCGCGATTCCACCGGATTCCTGGAGCTGGTTGAACATCC 1135
QY 260 TCATCCCTGTGTAGACCGGATTCGATATGTGAGAGTCTCAAGAAATTTCTATCAACG 319
DB 1134 TCATCCCTGTGTAGACCGGATTCGATATGTGAGAGTCTCAAGAAATTTCTATCAACG 1075
QY 320 TGCCTAGACATCGGCTGTGACTCTGACATGTAACTCTGCAATGANTGAGTCTTT 379
DB 1074 TGCCTAGACATCGGCTGTGACTCTGACATGTAACTCTGCAATGANTGAGTCTTT 1015
QY 380 ACCTGCGATCATGAGACCTTACAAAGCAGCTAAGGTGTGAGAGACCTGAGTGTGCG 439
DB 1014 ACCTGCGATCATGAGACCTTACAAAGCAGCTAAGGTGTGAGAGACCTGAGTGTGCG 955
QY 440 TCACCAGCTAGCTCAAAACACATGAGATGAGCTGGCAAACTCTGTGAGCAAG 499
DB 954 TCACCAGCTAGCTCAAAACACATGAGATGAGCTGGCAAACTCTGTGAGCAAG 895
QY 500 TCTTCGGGGAACGGGATCCCTGAATGCGACGATTTGTGATGCCATCAACCAAGCTGCTG 559
DB 894 TCTTCGGGGAACGGGATCCCTGAATGCGACGATTTGTGATGCCATCAACCAAGCTGCTG 835
QY 560 ACTGCTGGGATATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
DB 834 ACTGCTGGGATATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 775
QY 620 TGAAGAGTCTATGAGATGAGTGAAGCAGAGCGGCGGAAGGGCCACAGTTCTAG 679
DB 774 TGAAGAGTCTATGAGATGAGTGAAGCAGAGCGGCGGAAGGGCCACAGTTCTAG 715
QY 680 AGCTGAGGAGCCCGAGAGTGGCCATCATGTGCGCAGAGGAAGAAACAGGCCCCAGA 739
DB 714 AGCTGAGGAGCCCGAGAGTGGCCATCATGTGCGCAGAGGAAGAAACAGGCCCCAGA 655
QY 740 TCCTGGCTCCGACAGCAGAAAGGCTGAACAGATTAATTCAGCAGAGAGGCGCAGTG 799
DB 654 TCCTGGCTCCGACAGCAGAAAGGCTGAACAGATTAATTCAGCAGAGAGGCGCAGTG 595
QY 800 CAGTTCTGGCGAAGCCGAAGCTTAAGCTGAAGCTATTTCGATCCCTGGCTGACGCTGA 859
DB 594 CAGTTCTGGCGAAGCCGAAGCTTAAGCTGAAGCTATTTCGATCCCTGGCTGACGCTGA 535
QY 860 CACAACATATGAGATGAGATGAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCACGGCT 919
DB 534 CACAACATATGAGATGAGATGAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCACGGCT 475
QY 920 TCCTCAAACTGGCCGAAGAGCTCCACACTATCTACTGCTCCCAACCTGGCGATGTCA 979
DB 474 TCCTCAAACTGGCCGAAGAGCTCCACACTATCTACTGCTCCCAACCTGGCGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGCCAGTGCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGCCAGTGCAG 355
QY 1040 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 354 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 14
US-09-471-275-865/c
; Sequence 865, Application US/09471275
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US/09/471,275
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/235,076
; EARLIER FILING DATE: 1999-01-20

```
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 09/338,467
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/090,177
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 1022
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-384-941
```

```
Query Match      72.1%; Score 857; DB 31; Length 1384;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 140 GCCGCGCTCTCTGATGATGCCCCGAAACACCGTGTACTGTTGTCGCCGACGAGAGG 199
    |||
Db 138 gccgcgcctcctcgtgaltgcccgaaacacgctgctacgtctcgtgcgcgacgagag 197
OY 200 CCGGGGTGGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 259
    |||
Db 198 cctgggtgggagcgaaatggccgatccacggatcccgagcctgggttgacatcc 257
OY 260 TCATCCCTGTGTAGACCGGATCCGATGTGACAGCTCTCAAGGAATGTCTCAACG 319
    |||
Db 258 tcatacctcgtgtacgacggatccgataatgtcgaagatcccaaggaaatgtcacaag 317
OY 320 TGCCTGAGCAGTGGCTGTGACTCTGCACAATGTAACTCTGCAAAATGATGAGTCTCTT 379
    |||
Db 318 tgcctgagcagtcgctgtgactctgcacaatgttaactctgcaaatcgtatgagtcctt 377
OY 380 ACCGCGCATCATGAGGACCTTTACAGGCAAGTACGGTGGAGGAGACCTCGATGATCCG 439
    |||
Db 378 accgcgcacatcagaccccttaacaaagctacggctggagaccccgagatagcgg 437
OY 440 TCACCCAGTAGTCAAAACAACATGAGATCAGAGCTCGGCAACTCTCTNTGGACAAAG 499
    |||
Db 438 tcacccagctagctcaaaacaacatgagatcagagctcggcaaatctctcttgacaaag 497
OY 500 TCTTCGCGGGAACGGGAGTCCCTGTAATGCCAGATTGATGCCATTAACCAAGCTCTG 559
    |||
Db 498 tcttcgcggaacgggagtcctctgaatgccagctcttggaatgcatacaacagctctg 557
OY 560 ACTGCTGGGATATCCGCTGCTNGTATGATGATCAAGATATCAATGATGCAACCCCGGG 619
    |||
Db 558 actgctgggataatccgctcctcgtataagatcaagagatcccaatctctccacccggg 617
OY 620 TGAAGAAGTCTATGACAGATCAGGTGAGGAGCAGCGCGGAAACGGGCACAGTTCTAG 679
    |||
Db 618 tgaagaagctctatgcagatgcaggtgagagcgaggaacgggccaacagctctag 677
OY 680 AGCTGAGGGGACCCGAGATCGGCGCATCAATGTGGCAAGAAAGGAGAAACAGGCCAGA 739
    |||
Db 678 agctgagggagaccgagagctcgccatcaatgtgcagaaggaagaaacagggccaga 737
OY 740 TCTTGCGCTCGAAGCAGAAAGGCTGAAGATTAATTCAGGCGAGCAGAGAGGCCAGTG 799
    |||
Db 738 tcttgcgctcgcgaagcagaaagctgaacagataaaccagcgaggaagggccagtg 797
OY 800 CAGTTCTGGGAAGGCCAAGCTAAAGCTAAGCTATTTCGAATCCCTGGCTGACGCTTGA 859
    |||
Db 798 cagttctgggaagggcgaagcctaagctatctgaatccctggcgcgagctctga 857
OY 860 CACAACATATGAGATGACGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAAGCGCT 919
    |||
Db 858 caaacaataatgagatgacgagcttcactgactgtggccgagcagtatgtcaagcgct 917
OY 920 TCCGCAAGTGGCCGAAGACTTCAACACTATCTACTAGCCCTCAACCCCGGGAGTCTCA 979
    |||
Db 918 tcgcgaagtggccgaagacttcaaacatactctactcgtccctccacccctggagtgta 977
```

```
OY 980 CCAGCATGTGGCTCAGGCCATGGGTATATGGAGCCCTCACCAAGGCCCATGGCCAG 1039
    |||
Db 978 ccagcatggtggctcagggcatgggtatataatgagccctcaccaagccccagtgccag 1037
OY 1040 GGACTCCAGACTCAGTCTCTCAGTGGAGAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
    |||
Db 1038 ggaactccagactcagctctcagtgagagcagagagtgctccagggatcagatgcaagt 1096
```

RESULT 12

```
PCT-US01-03800A-1001/c
; Sequence 1001, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1001
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-03800A-1001
```

```
Query Match      72.1%; Score 857; DB 1; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 140 GCCGCGCTCTCTGATGATGCCCCGAAACACCGTGTACTGTTGTCGCCGACGAGAGG 199
    |||
Db 1254 GCCGCGCTCTCTGATGATGCCCCGAAACACCGTGTACTGTTGTCGCCGACGAGAGG 1195
OY 200 CCGGGGTGGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 259
    |||
Db 1194 cctgggtgggagcgaaatggccgatccacggatcccgagcctgggttgaaatcc 1135
OY 260 TCATCCCTGTGTAGACCGGATCCGATGTGACAGAGTCTCAAGGAATGTGATCAACG 319
    |||
Db 1134 tcatacctcgtgttagacccgatccgatatgtgcagagcttcaaggaaatgtcatcacg 1075
OY 320 TGCCTGAGCAGTGGCTGTGATGATGCAAAATGTAATGTAATGATGAGTCTCTT 379
    |||
Db 1074 tgcctgagcagtggtgctgtgactctgcacaatgttaactctgcaaatgcagatgagtccttt 1015
OY 380 ACCTGCGCATCATGAGACCCCTTACAAAGCAAGCTACGGTGTGGAGAGACCTGATATGCC 439
    |||
Db 1014 acctgcgcatcatgagaccccttacaaagcagctacggtgtggagagacccctgatatggccg 955
OY 440 TCACCCAGTAGTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
    |||
Db 954 tcacccagtagtcaaaacaacatgagatcagagctcgccaaactctcttgacaaag 895
OY 500 TCTTCGCGGGAACGGGAGTCCCTGAATGCCAGATTGTGATGTCATCAACCAAGCTGCTG 559
    |||
Db 894 tcttcgcggaacgggagtcctcctaagtcagcatgttgatgacatcaacaaagctgctg 835
OY 560 ACTGCTGGGATATCCGCTGCTNCGTTATGAGATCAAGGATATTCATGTGTCACCCCGGG 619
    |||
Db 834 actgctgggataatccgctcctcgttatgagatcaaggatattccatgtgcccacccggg 775
OY 620 TGAAGAAGTCTATGACAGATCAGGTGAGGAGCAGAGCGCGGAAACGGGCACAGTTCTAG 679
    |||
Db 774 tgaagaagctctatgacagatcaggtgagagcagagcgcggaacgggcacacagttctag 715
```

QY 440 TCACCCAGCTAGCTCAAAACCATGAGATCAGAGCTCGGCAAACTCTGTTGACAAAG 499
 Db 438 tcaccacagctagctcaacaacccatgagatcagagctcgcgaactctctctgacaaaag 497
 QY 500 TCTTCCGGGAACGGAGATCCCTGTAATGCCAGCATTTGTGATGCCATCAACCAACTCTG 559
 Db 498 tcttcgaggaacggagatccctgtaattgcagcatctgtgagatccaaacacagctgctg 557
 QY 560 ACTGCTGGGGATCCGGCTGCTGCTGTTATGATCAAGGATATCATGTGCCACCCGGG 619
 Db 558 acgctgaggatccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 617
 QY 620 TGAAGAGTCTATGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
 Db 618 tgaagagctatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 677
 QY 680 AGTCTGAGGGAG 739
 Db 678 agctgagaggag 737
 QY 740 TCTGAGCTCCGAAAG 799
 Db 738 tcctgagctccgaaag 797
 QY 800 CAGTTCTGGCAAGGAG 859
 Db 798 cagttctggcgaag 857
 QY 860 CACAACATATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 919
 Db 858 cacaacataatgagagatgagagatgagagatgagagatgagagatgagagatgagagatg 917
 QY 920 TCTCCAACTGGCCAAAG 979
 Db 918 tctccaaactggccaaag 977
 QY 980 CCAGCATGGTGGCTCAGGCGCATGGGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1039
 Db 978 ccagcatgggtggctcaggcgcatgggtgtatgagagagagagagagagagagagagagag 1037
 QY 1040 GGAGTCACAGACTCCTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 Db 1038 ggagtcacagactcctccatgagagagagagagagagagagagagagagagagagagag 1096
 RESULT 10
 US-09-652-914-7997
 ; Sequence 7997, Application US/09652914
 ; GENERAL INFORMATION:
 ; APPLICANT: Shyjan, Andrew W.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.1193-001
 ; CURRENT APPLICATION NUMBER: US/09/652,914
 ; PRIOR FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 60/152,112
 ; NUMBER OF SEQ ID NOS: 9677
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7997
 ; LENGTH: 1384
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-652-914-7997

Query Match 72.1%; Score 857; DB 25; Length 1384;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 140 GCGCGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 199
 Db 140 GCGCGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 199

Db 138 gccgcgcctcctctgagatgccccgaacacgctggtactctctgctgagcagagagag 197
 QY 200 CCTGGGTGGTGGAGCGCAATGGCGGATTCACACCGAGTCCGAGCCGTTGATTAACATCC 259
 Db 198 cctgggtgggtggagcgaatggcggttcacacccgagttccgagccgttggatgaacatcc 257
 QY 260 TCAATCCCTGTGTAGACCGGATCCGATATGTGCAAGATCTCAAGAAATTTGTATCAACG 319
 Db 258 tcaatccctgtgttagaccggatccgatattgtgcaagatctcaagaaatttgtatcaacg 317
 QY 320 TGCCGTGAGATGCGGTGTGATCTGACATATGTAATCTGCAATGATGATGATGATGATGAT 379
 Db 318 tgccgtgagatgcggtgtgattctgacatattgttaattctgcaattgataatgataatgata 377
 QY 380 ACTGCGCATCATGAGACCCCTTACAGAGCAAGCTACGCTGAGAGAGAGAGAGAGAGAGAG 439
 Db 378 actgcgcatcatgagaccccttaccagagcactacagagagagagagagagagagagagagag 437
 QY 440 TCACCCAGCTAGCTCAAAACCATGAGATCAGAGCTCGGCAAACTCTGTTGACAAAG 499
 Db 438 tcaccacagctagctcaacaacccatgagatcagagctcgcgaactctctctgacaaaag 497
 QY 500 TCTTCCGGGAACGGAGATCCCTGTAATGCCAGCATTTGTGATGCCATCAACCAACTCTG 559
 Db 498 tcttcgaggaacggagatccctgtaattgcagcatctgtgagatccaaacacagctgctg 557
 QY 560 ACTGCTGGGGATCCGGCTGCTGCTGTTATGATCAAGGATATCATGTGCCACCCGGG 619
 Db 558 acgctgaggatccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 617
 QY 620 TGAAGAGTCTATGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
 Db 618 tgaagagctatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 677
 QY 680 AGTCTGAGGGAG 739
 Db 678 agctgagaggag 737
 QY 740 TCTGAGCTCCGAAAG 799
 Db 738 tcctgagctccgaaag 797
 QY 800 CAGTTCTGGCAAGGAG 859
 Db 798 cagttctggcgaag 857
 QY 860 CACAACATATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 919
 Db 858 cacaacataatgagagatgagagatgagagatgagagatgagagatgagagatgagagatg 917
 QY 920 TCTCCAACTGGCCAAAG 979
 Db 918 tctccaaactggccaaag 977
 QY 980 CCAGCATGGTGGCTCAGGCGCATGGGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1039
 Db 978 ccagcatgggtggctcaggcgcatgggtgtatgagagagagagagagagagagagagagag 1037
 QY 1040 GGAGTCACAGACTCCTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 Db 1038 ggagtcacagactcctccatgagagagagagagagagagagagagagagagagagagag 1096
 RESULT 11
 US-09-808-384-941
 ; Sequence 941, Application US/09808384
 ; GENERAL INFORMATION:
 ; APPLICANT: Geating, David P.
 ; APPLICANT: Holtzman, Douglas A.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES DERIVED FROM A
 ; FILE REFERENCE: 1600.1025-002
 ; CURRENT APPLICATION NUMBER: US/09/808,384


```
|||||
Db 918 tctccaacccgccaaggactccaacactctcactcgcctcccaaccccgagtgca 977
OY 980 CCAGCATGTGTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAGAGCCCGCAGTGCAG 1039
Db 978 ccagcatggtggtctcagggccatgggtatataatgagccctccacaagcccccagtgccag 1037
OY 1040 GGACTCCAGACTCCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagtgaggagcagagatgtccagggtacagatgcaagt 1096
```

RESULT 8

```
US-09-652-816-7419
; Sequence 7419, Application US/09652816
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1177-001
; CURRENT APPLICATION NUMBER: US/09/652,816
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,111
; NUMBER OF SEQ ID NOS: 9647
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7419
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-816-7419
```

```
Query Match 72.1%; Score 857; DB 25; Length 1384;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 140 GCCGCGCTCTCTGTGATTGCCCCGAAACACCGTGTACTGTTGTCGCCGACAGAGAG 199
Db 138 gccgcgctctctgtgattgccccgaacaacgltgtactgttcgtgcgcagcagagag 197
OY 200 CCGGGTGTGGAGCCGAATGGCGGATTCACCGGATCCTGGAGCCGTGTTGAACATCC 259
Db 198 cctgggtgtggagccgaatggcggaattccacggatccctggagccgtgttgaacatcc 257
OY 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAAGAAATGTGTATCAACG 319
Db 258 tcattccctgtgttagaccggatccgatatgtgcagagttccaaagaaattgtcatcaag 317
OY 320 TGCCTGACAGTGGCTGTGACTCTCGACAAATGTAATCTGCAAAATCGATGAGTCCCTTT 379
Db 318 tgcctgagagtgctgactctcgacaatgttaactctgcaaatcgatgagatcctt 377
OY 380 ACCTGGCATGAGGACCCCTTACAAGCAAGTACAGTGTGGAGAGACCCGATATGCCC 439
Db 378 acctggcatgaggaaccttacaagcaagctacgltgaggagaccctgagtalagccg 437
OY 440 TCACCCAGTAGCTCAACAAACCATGAGATCAGAGCTCGGCAAACTCTTNTGACAAAG 499
Db 438 tcaccagctagctcaaaaacaacatagatcagagctcggcaaacctctcttgacaag 497
OY 500 TCTTCGGGAGACGGAGTCCCTGATGACGACATTTGTGATGCCANTCAACAAAGTCTTG 559
Db 498 tcttcgggagacggagtcctctgacgacatctgtagatgacatcaaaccaagctgctg 557
OY 560 ACTGCGGGGTATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
Db 558 actgctggggtatccgctcgttcgtttagagatcaagataltcattgtgcccccgg 617
OY 620 TGAAGAGTCTAGTACAGATGAGAGGAGGAGGCGGAAACGGCCACAGTCTTAG 679
Db 618 tgaagagctctatgacagatgagagtgaggcagagcggcggaacgggcccacagttctag 677
```

```
OY 680 AGTCTGAGGGAGCCCGAGAGTCCGGCATCATATGTGCGAGGAAGGAAACAGGCCAG 739
Db 678 agtctgagggagcccgagagtgccatcaatgttgagagaaggagaaacagggccaga 737
OY 740 TCTGTGCTTCGGAAGCAGAAAGGCTTCAACAGATTAATTCAGGACGAGGAGCCAGTGC 799
Db 738 tctgtgcttcggaagcagaaaggcttcaacagataaaltcagcagcagagagccagtg 797
OY 800 CAGTTTGGGAGAGGCCAAGGCTTAAAGCTTAAGCTTATTCGAATTCCTGGCTGACGCTCTGA 859
Db 798 cagttctgggagagggccaagctcaagctgaagctatcgaaatccgtggtccgctctga 857
OY 860 CACAACATTAATGAGATGACAGACGCTTCACTGACTGTGGCCGAGCATGTACAGCCGT 919
Db 858 cacaacataatgagatgacagacttcaactgactgtggtgcgagcatgtacgagcgt 917
OY 920 TCTTCAACTGGCCGACGAGTCCCAACACTTCTACTGCCCCCTCAACCTTGGGATGTCA 979
Db 918 tcttcaaacctggccaaggactccaacactatcctactgcctccaacccctgagatgtca 977
OY 980 CCAGCATGTGTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAAGCCCGAGTGCAG 1039
Db 978 ccagcatggtggtctcagggccatgggtatataatgagccctccacaagcccgagtgccag 1037
OY 1040 GGACTCCAGACTCCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagtgaggagcagagatgtccagggtacagatgcaagt 1096
```

RESULT 9

```
US-09-652-913-8809
; Sequence 8809, Application US/09652913
; GENERAL INFORMATION:
; APPLICANT: Fald, Dean R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1174-001
; CURRENT APPLICATION NUMBER: US/09/652,913
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,107
; NUMBER OF SEQ ID NOS: 10833
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8809
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-913-8809
```

```
Query Match 72.1%; Score 857; DB 25; Length 1384;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 140 GCCGCGCTCTCTGTGATTGCCCCGAAACACCGTGTACTGTTGTCGCCGACAGAGAG 199
Db 138 gccgcgctctctgtgattgccccgaacaacgltgtactgttcgtgcgcagcagagag 197
OY 200 CCGGGTGTGGAGCCGAATGGCGGATTCACCGGATCCTGGAGCCGTGTTGAACATCC 259
Db 198 cctgggtgtggagccgaatggcggaattccacggatccctggagccgtgttgaacatcc 257
OY 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAAGAAATGTGTATCAACG 319
Db 258 tcattccctgtgttagaccggatccgatatgtgcagagttccaaagaaattgtcatcaag 317
OY 320 TGCCTGACAGTGGCTGTGACTCTCGACAAATGTAATCTGCAAAATCGATGAGTCCCTTT 379
Db 318 tgcctgagagtgctgactctcgacaatgttaactctgcaaatcgatgagatcctt 377
OY 380 ACCTGGCATGAGGACCCCTTACAAGCAAGTACAGTGTGGAGAGACCCGATATGCCC 439
Db 378 acctggcatgaggaaccttacaagcaagctacgltgaggagaccctgagtalagccg 437
```


Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
140 GCCGGCCCTCTGGATGCCCCGAAACAGCGTGTACTGTCTGCGCCGACGAGAG 199
    |||||||
138 gccgcgccccctctgattgccccgaacaacgctgtaactgtctgccccgacgagag 197
    |||||||
200 CCTGGGTGTGAGGCAATGAGGCCATTCCACCGGATCCTGAGCCTGTGTAACATCC 259
    |||||||
198 cctgggtgtgtgaggaataggccgattccaccggaatccttgagcctgtgttgaaatcc 257
    |||||||
260 TCATCCCTGTGTAGACCGGATCCCATATGTGCAGAGTCTCAGGAAATTGTCATCAGC 319
    |||||||
258 tcattccctgtgttagaccggaatcgaatgtgcagagctcgaagaaattgtcataacg 317
    |||||||
320 TCCGTGAGAGTGGCTGTGACTCTGCAGTAATGTAATCTGCAATCGATGGAGTCCCTT 379
    |||||||
318 tgcctgagagcagctcgctgtgactctgcgaataatgtaactctgcgaatcgatgagctctt 377
    |||||||
380 ACCGTGCGCATCATGAGACCTTACAAGGCAAGCTAGCGTGTGAGAGACCTGTGATGCCG 439
    |||||||
378 accgtgcatactatgagaccttaacaagctacggtgtgagagacctgtgtatgccc 437
    |||||||
440 TCACCCAGCTACTCAACACACCATGATGAGTCAAGCTCGGCAACTCTTNTGACCAAG 499
    |||||||
438 tcaccagactaagctcaacaacacatgagatcagagctcgcaaaactctctcgagcaag 497
    |||||||
500 TCTTCCGGGAAGGGAGAGTCCCTGTAATGCCAGATGTGATGCCATCAACAGAGTCTG 559
    |||||||
498 tcttcgggaagaggagatccctggaatctgagcatgtgagatgcataacacagctgcg 557
    |||||||
560 ACTGTGGGGTATCCGCTCCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCGGG 619
    |||||||
558 actgtgggggtatccgctcctcgttatgagatcaagataatccatgtgccaacccgg 617
    |||||||
620 TCAAGAGCTATGCAAGATGAGTGTGAGGACGAGCGGGGAAAGGGCCACATGTTCTAG 679
    |||||||
618 tgaagagactatgacatgacagctgagagcgagcgagcgagaaacgggccaagtctag 677
    |||||||
680 ACTGTGAGGAGACCGGAGAGTGGCCATCAATGTGCGACAGAGGAAAGAGGCCAGAG 739
    |||||||
678 acttcgagggagcccgagagctcgcatcaatgttggcagaagaggaaacgggccagaa 737
    |||||||
740 TCCCTGGCTCCGACAGCAAAAAGCTGAACAGATAATCAGCGACGAGAGGCCAGTG 799
    |||||||
738 tccctggcctcgaagcagaagaaagctgaaacagataatcagagaaagagagccagtg 797
    |||||||
800 CAGTTCCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGGAATCCTGGCTGACGCTTGA 859
    |||||||
798 cagttctggcgaaagccaaagctlaaagctgaagctatctgaaatcctggtcgaagctga 857
    |||||||
860 CACAACATATGAGATGAGAGAGCTTCACTGACTGTGCGCAGAGAGTATGTCAGCGCT 919
    |||||||
858 caacaataatgagatgagatgagcagctcactgaatgttggcagagctatgtcagagcgct 917
    |||||||
920 TCTCCAACTGGCCCAAGAGCTCCAAACATATCTACTGCGCTCCAAACCTGTGCGATGTCA 979
    |||||||
918 tctccaaactggcccaagagctccaacatactactgcccctccaacccctgtgagatgca 977
    |||||||
980 CCAGCATGTGTGCTCAGGCGCATGGGTATATGAGAGCCTTCACCAAAAGGCCAGTCCAG 1039
    |||||||
978 ccagcatgtgtgtcagcagctcagtgatatactgagagccctcacaagagcccaagtgccag 1037
    |||||||
1040 GGAATCCAGATCTACTCTCAGTGGAGAGCAGAGAGATGTCAGGGTATAGATGCAAGT 1098
    |||||||
1038 ggaatccagatctactctcagtgaggagcagcagagatgtccagggatcacagatgtcaagt 1096
    |||||||
```

RESULT 7
US-09-652-127-7358
: Sequence 7358, Application US/09652127
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.

```
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1183-001  
; CURRENT APPLICATION NUMBER: US/09/652,127  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/151,134  
; NUMBER OF SEQ ID NOS: 10475  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 7358  
; LENGTH: 1384  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-652-127-7358
```

Query Match 72.1%; Score 857; DB 25; Length 1384;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
140 GCCGGCCCTCTGGATGCCCCGAAACAGCGTGTACTGTCTGCGCCGACGAGAG 199
    |||||||
138 gccgcgccccctctgattgccccgaacaacgctgtaactgtctgccccgacgagag 197
    |||||||
200 CCTGGGTGTGAGGCAATGAGGCCATTCCACCGGATCCTGAGCCTGTGTAACATCC 259
    |||||||
198 cctgggtgtgtgaggaataggccgattccaccggaatccttgagcctgtgttgaaatcc 257
    |||||||
260 TCATCCCTGTGTAGACCGGATCCCATATGTGCAGAGTCTCAGGAAATTGTCATCAGC 319
    |||||||
258 tcattccctgtgttagaccggaatcgaatgtgcagagctcgaagaaattgtcataacg 317
    |||||||
320 TCCGTGAGAGTGGCTGTGACTCTGCAGTAATGTAATCTGCAATCGATGGAGTCCCTT 379
    |||||||
318 tgcctgagagcagctcgctgtgactctgcgaataatgtaactctgcgaatcgatgagctctt 377
    |||||||
380 ACCGTGCGCATCATGAGACCTTACAAGGCAAGCTAGCGTGTGAGAGACCTGTGATGCCG 439
    |||||||
378 accgtgcatactatgagaccttaacaagcagatcaggtgtgagagacctgtgtatgccc 437
    |||||||
440 TCACCCAGCTACTCAACACACCATGATGAGTCAAGCTCGGCAACTCTTNTGACCAAG 499
    |||||||
438 tcaccagactaagctcaacaacacatgagatcagagctcgcaaaactctctcgagcaag 497
    |||||||
500 TCTTCCGGGAAGGGAGAGTCCCTGTAATGCCAGATGTGATGCCATCAACAGAGTCTG 559
    |||||||
498 tcttcgggaagaggagatccctggaatctgagcatgttggcagaagaggaaacgggccagaa 557
    |||||||
560 ACTGTGGGGTATCCGCTCCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCGGG 619
    |||||||
558 actgtgggggtatccgctcctcgttatgagatcaagataatccatgtgccaacccgg 617
    |||||||
620 TGAAGAGCTATGCAAGATGAGTGTGAGGACGAGCGGGGAAAGGGCCACAGTTCCTAG 679
    |||||||
618 tgaagagactatgacatgacagctgagagcgagcgagcgagaaacgggccaagtctag 677
    |||||||
680 ACTGTGAGGAGACCGGAGAGTGGCCATCAATGTGCGACAGAGGAAAGAGGCCAGAG 739
    |||||||
678 acttcgagggagcccgagagctcgcatcaatgttggcagaagaggaaacgggccagaa 737
    |||||||
740 TCCCTGGCTCCGACAGCAAAAAGCTGAACAGATAATCAGCGACGAGAGGCCAGTG 799
    |||||||
738 tccctggcctcgaagcagaagaaagctgaaacagataatcagagaaagagagccagtg 797
    |||||||
800 CAGTTCCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGGAATCCTGGCTGACGCTTGA 859
    |||||||
798 cagttctggcgaaagccaaagctlaaagctgaagctatctgaaatcctggtcgaagctga 857
    |||||||
860 CACAACATATGAGATGAGAGAGCTTCACTGACTGTGCGCAGAGAGTATGTCAGCGCT 919
    |||||||
858 caacaataatgagatgagatgagcagctcactgaatgttggcagagctatgtcagagcgct 917
    |||||||
920 TCTCCAACTGGCCCAAGAGCTCCAAACATATCTACTGCGCTCCAAACCTGTGCGATGTCA 979
    |||||||
```

```

Db 618 tgaagaagctctacgagatcaggtggaagcagcgccgaaacgagccacagttctag 677
Oy 680 AGACTGAGGGAGCCCGAGAGTCCGACATCATGTGCGAGAGGAACAAACAGGCCAGAG 739
Db 678 agctctgagggagaccgagagagtcggtccatcctcagtgagagaaagaaacagccagag 737
Oy 740 TCTTGCCCTCCGAGCAGAAAGAGCTGACAGATTAATTCAGGAGCAGAGAGGCCAGTG 799
Db 738 tcttgccctcgaagcagaaagctgaacagtaatacgaagcagcgagagccaggtg 797
Oy 800 CAATTCTGGGGAAGGCCAAGCTTAAGCTAAGCTATTGCAATTCCTGCTGACGCTTGA 859
Db 798 cagttctgaggaagcgaagcgaagctgaagctatcgaatccctgctgcagctctga 857
Oy 860 CACAACATATGAGATGACGAGACTTCAGTCTGAGTGGCCGAGCAATATGTCAGCGCGT 919
Db 858 cacaacataatgagatgacgagctcactgactgagcagcagatcagtcagcggt 917
Oy 920 TCTCCAACTGGCCCAAGAGACTCCACACTATCTACTGCCCCCAACCCCTGGCGATGCA 979
Db 918 tctccaactgcccgaagactccacaactatctcactgcccctccaaacctgscagtca 977
Oy 980 CCAGCATGCTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAGAGCCCGAGTCCAG 1039
Db 978 ccagcatggtggtcagcgccatggtgtatagagccctcaccaagccccagtgccag 1037
Oy 1040 GGACTCAGACTCACTCTCCAGTGGAGCAGAGATGTCAGAGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactcctcagtggaagcagagagatgcccaggtacagatgcaagt 1096

```

RESULT 5

```

US-09-162-7309
; Sequence 7309, Application US/09649162
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Richardson, Jennifer
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1181-001
; CURRENT APPLICATION NUMBER: US/09/649,162
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,057
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9990
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7309
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-162-7309

```

Query Match 72.1%; Score 857; DB 25; Length 1384;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 140 GCCGGCCCTCTTGATGATGCCCGAAGACGCTGATGTTGTCGCCGACGAGAGG 199
Db 138 gccgcgctctctctgattgctcccgaaacacgctggtactctgctccgacagagag 197
Oy 200 CCGGGGTGGGAGCAATGGGCGGATTCACGCGATCCGAGAGCCCTGGTTTAACATCC 259
Db 198 cctgggtggtgagcgaatgagcgaatccacggaatcccgagcctgggttgaacatcc 257
Oy 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAGGAATGTGATCAAG 319
Db 258 tcatccctgtgttagcgcggaatcgatagtcgagagctcccaagaaatgtcatcaacg 317
Oy 320 TGCCTGAGCAGTGGCTGTGACTCTCGACAAATGAACTGATGAGAGTCTTT 379

```

```

Db 318 tgcctgagcagtcggtctgactctcgacaatgtaactctgcaaatcgatgagctctt 377
Oy 380 ACCTGGCATCATGAGACCCCTTACAGCAAGCTACGGTGTGGAGAACCTGAGATGCGG 439
Db 378 acctgcatcatalgagacccttacaaagcgaagctacggtggtgagagccgagatagcgg 437
Oy 440 TCACCCAGCTAGCTCAACACCATGAGATCAGAGCTCGGCAACCTCTTNTGACAAAG 499
Db 438 tcaaccgctagcttcaaaacacatgagatcagagctccgccaactctctcigacaag 497
Oy 500 TCTTCCGGGAAGGAGAGTCCCTGAATGCCAGCATTTGGATGCCATACCAAGCTGCTG 559
Db 498 tcttcggaacggaagctccctgaatgcccagcatltgtga tgcacataacaaagctgtctg 557
Oy 560 ACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
Db 558 actgctgggtatccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 617
Oy 620 TGAAGAGTCTATGACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679
Db 618 tgaagaagctctacgagatgcaagtggaagcagagcgagaaacgagccacagttctag 677
Oy 680 AGCTGAGGGAGCCCGAGAGTCCGCTCATATGTGCGAAGAGGAAGAAACAGGCCAGAG 739
Db 678 agctcagggagaccgagagctcgccatcctcagtggtgcaagaaggaagaaacagccaga 737
Oy 740 TCTTGCCCTCCGAGCAGAAAGAGCTGAACAGATTAATTCAGGAGCAGAGAGGCCAGTG 799
Db 738 tcttgccctcgaagcagaaagctcgaacagatgaatccagcagagagagagccaggtg 797
Oy 800 CAATTCTGGGGAAGGCCAAGGCTAAAGCTTAAGCTATTCGAATCTTGCTGACGCTTGA 859
Db 798 cagttctgaggaagcgaagcgaagctgaagctatcgaatccctgagcagcagctctga 857
Oy 860 CACAACATATGAGATGACGAGCTTCACTGACTGTGGCCGAGCAGATGTCAGCGCGT 919
Db 858 cacaacataatgagatgacgagctcactgactggtggtgagcaggtatgtaagcggt 917
Oy 920 TCTCCAACTGGCCCAAGAGACTCCACACTATCTACTGCCCCCAACCCCTGGCGATGCA 979
Db 918 tctccaactgcccgaagactccacaactatctcactgcccctccaaacctgscagtca 977
Oy 980 CCAGCATGCTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAGAGCCCGAGTCCAG 1039
Db 978 ccagcatggtggtcagcgccatggtgtatagagccctcaccaagccccagtgccag 1037
Oy 1040 GGACTCAGACTCACTCTCCAGTGGAGCAGAGATGTCAGAGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactcctcagtggaagcagagagatgcccaggtacagatgcaagt 1096

```

RESULT 6

```

US-09-652-109-7925
; Sequence 7925, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7925
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-7925

```

Query Match 72.1%; Score 857; DB 25; Length 1384;

```

1  TITLE OF INVENTION: HUMAN PROSTATE STROMAL LIBRARY
2
3  FILE REFERENCE: MLN98-25p
4
5  CURRENT APPLICATION NUMBER: US/09/338,467
6
7  CURRENT FILING DATE: 1999-06-22
8
9  EARLIER APPLICATION NUMBER: 60/090,177
10
11 EARLIER FILING DATE: 1998-06-22
12
13 NUMBER OF SEQ ID NOS: 1022
14
15 SOFTWARE: FastSeq for Windows Version 3.0
16
17 SEQ ID NO 941
18
19 LENGTH: 1384
20
21 TYPE: DNA
22
23 ORGANISM: Homo sapiens
24
25 US-09-338-467-941

```

Query Match	72.1%;	Score 857;	DB 17;	Length 1384;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 957;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

Oy	140	GC	CGCGCTCTCT	CGGATGCGCGCCGGA	AAACCGCGGATCGTTCTGCGCCAC	CAGAGAG	199	
Db	138	gc	cgcgctctct	cttgatctgcccga	aacacgctgtaactcttgccgca	gaggaag	197	
Oy	200	CCT	GAGTGTGTG	AGCAGATGAGGCGCGATTC	CAACCGAGTCTCGAGACCTG	GTTTGAATCC	259	
Db	198	cct	ggtgtgtg	agcgaatgggcccgat	tccaacggatccctggagcctggttt	baaatcc	257	
Oy	260	TC	ATCCCTGTGTT	AGACCGGATCCGATATGTG	CAGAGTCTCAAGGAATTTG	TATCAACG	319	
Db	258	tca	tccctgtgt	agaccggatccgatatgtgcag	gtctccaaggaattgtatcaacg		317	
Oy	320	TG	CTTGAGACAGT	CGCGCTGTGACTTTCGACA	TGTAACTCTGCAATTCGATG	AGAGTCTTT	379	
Db	318	tg	ccgtgagcag	ctcgctgtgactcttcga	acaatgttaactctgcaatctgat	tgatccctt	377	
Oy	380	AC	CTCGCATCAT	GAGACCCCTTACAAGCAAG	CTCGGTGTGGAGGACCTTC	AGATGCGG	439	
Db	378	a	ctgcgcatact	gagaccccttaacaggcaac	ctacggttgtagagacctgtatg	ccg	437	
Oy	440	TC	ACCACACTAG	CTCAAAACACATGAGATAG	AGCTCGGCAAACTCTCTNTG	AGCAAG	499	
Db	438	tca	cccgactgtac	ctcaaacacacatg	atgatacgctccggaacatctctct	tgacaag	497	
Oy	500	TC	TTTCGGGAA	CGGAGTCCCTGATGCGAC	ATTTGTGATGCCATCAAC	CAAGCTGCTG	559	
Db	498	tct	tcgcggaac	cggaagtccctctgatactgc	agcatctgtgtatgcatcaac	aaagctgtcg	557	
Oy	560	AC	TGTTGGGGAT	ATCCGCTGCTNGCTTATG	AGATATTCATGTGCCAC	CCCCGGG	619	
Db	558	a	ctgtctgggtat	ctccgctcgtctatgatacca	agatatctcatctgtgcc	ccccggg	617	
Oy	620	TG	AAAGACTCTAT	GCAGATGCAGGTGGAGG	CAGCGGCAAA	CGGGCCACAGTTCTAG	679	
Db	618	tga	aaagatctat	gcagctgcaagtlgtagagga	agagcgcggaacggtcc	acagcttctag	677	
Oy	680	AG	CTGAGAGGG	AGCCGAGAGTCCGACAT	ATGTTGGCAGAGGAAG	GAAMAACAGGCCAGA	739	
Db	678	ag	tctgagggag	cccgagagctcgccatcaat	ctgtgcgaagggaaataac	caagccaga	737	
Oy	740	TC	CTTGCGCT	CCGAGCAGAAAGCTGA	CAGATPAATCAGG	CAGCAGGAGAGGCCAGTG	799	
Db	738	tcc	tgcctctcg	caagcagaataagctg	ataacataatcaagcag	agaggtccagtg	797	
Oy	800	CAG	TTCTGGCCGA	AGGCCCAAGGCTTAAAG	CTGAAGCTATTCGAATCT	CGGTGCGACGCTCTGA	859	
Db	798	ca	gttctctgc	gaggaagccaaagctlaa	agcttgaagctatctgcaat	cttcggtctgcgacttga	857	
Oy	860	CAC	ACATPAAT	TGAGATGCACAGCTT	CACGACTTACGAGCTGT	GCGCAGCAGATATGTCAG	919	
Db	858	cac	aacataat	gaggaatgcagagcttca	ctcgactgtgtgcccga	gcagtatgtccagcgct	917	
Oy	920	TCT	CCAACTGG	CCCAAGACTTC	CAACACTATCTACTG	CCCTCCAA	CCCTGGCGATGTCA	979

Db	918	tctccaacatcigcccaagcagcattccactactccctcccaacccctggcgatgca	977
Qy	980	CCAGCATGTTGGCTCAGGCCATGGGTATTATATGAGCCCTCACCACCAAGCCCATGTCGAG	1033
Db	978	ccagaatcvtgctcgaagccatctggatctatctgtagccctccacaagcccctcgtccag	1033
Qy	1040	GGACTCCAGACTCACTCTCCAGTGGGAGCAGACAGATGTCACAGGTTACAGATGCAAGT	1098
Db	1038	ggaccctcagactcaatctccctcagctggagcagcaagatctccaggggtatcagatgcagat	1096

RESULT 4
US-09-644-869-8024

```

; Sequence 8024, Application US/096444869
; GENERAL INFORMATION:

```

```

; APPLICANT: Shyjan, Andrew W.
; APPLICANT: McCarthy, Sean A.

```

APPLICANT: Holtzman, Douglas A.
APPLICANT: Monahan, John

;; APPLICANT: Richardson, Jennifer
;; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

```

; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1182-001

```

;; CURRENT APPLICATION NUMBER: US/09/644
;; CURRENT FILING DATE: 2000-08-28

;; PRIOR APPLICATION NUMBER: 60/151,062
;; PRIOR FILING DATE: 1999-08-27

```
; NUMBER OF SEQ ID NOS: 9708
; SOFTWARE: FastSeq for windows Version 4.
```

```

; SEQ ID NO 8024
; LENGTH: 1384
;

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

05-09-644-869-8024

Query Match	72.1%;	Score 857;	DB 25;	Length 1384;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 957; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	140	GGCGGCGCCTCTTGATGGTCCCGGAAACACCGTGATCTGTTGTGTCGCGCAGCAGAGG	199
Db	138	ggcgcgccctccctctgattgcgcccgaaacacacgylgtactcttgctgcgcagcagagga	197
QY	200	CGTGGTGTGTGGAGGAATGGGGCGATTCCACGGGATTCGTGGACCGGTTTGAAATCC	259
Db	198	ccctggagtcgtgagagcaatcgccgattccacccggatcccgagaccggtttgaacatcc	257
QY	260	TGATCCCTGTGTTAGACCGGATTCGGATTGTGCAGAGTCTTCAAGAAATTTGTCATCAG	319
Db	258	tcatccctgtgttagaccggatccgattgtgcagagtlctcaagaaatgtctcaacg	317
QY	320	TGCGTGAGACGCGGCGCTGTGACTCTCGCATGTGAATCACTGCAATGAGTGGATCCTTT	379
Db	318	tgccctgagagcgcgctgtgactcttcgcgaacttgaacctcgcgaatactcgaatggatccctc	377
QY	380	ACCTGGCATCATGAGACCTTACAAAGGCACATCGGTGTGGAGAGACCTTGATATCCG	439
Db	378	acctgagcatatgacccttacaaggaagctacggtgtgagagaccctgagatgcg	437
QY	440	TGACCCAGTACTACTAAACACCATGAGATATAGAGCTCGGAACTCTGTTGGACAAAG	499
Db	438	tcaaccaagctagctcaaaacaacatgagatcagagctcggaanaactctctgtgacaag	497
QY	500	TCTTTCGGGAACGGGAGTCCCTGGAATGGCACATTGTGTGGATCCATCAACCAAGTCTG	559
Db	498	tcttcgcggaacagcgagctccctgaaatgcgcagcatgtgtagtgcatacaacaagctgtc	557
QY	560	ACTGCTGGGGGATTCGGCTGCTGCTGCTGCTTATGAGATCAAGATATCATGTGCAACCCGGG	619
Db	558	actgctcgaggatctccgctgcgcctcgcttatgagatcaagatattccatggtccaaaccggg	617
QY	620	TGAAAGAGCTATGCAGATGCAGGTGAGGAGAGACCGCGGAAACGGGCGCACAGTTTAC	679

;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 550
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 566
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 606
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 607
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 608
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 609
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 610
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/049, 611
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/050, 901
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/052, 989
;; PRIOR FILING DATE: 1997-06-13
;; PRIOR APPLICATION NUMBER: 60/051, 919
;; PRIOR FILING DATE: 1997-07-08
;; PRIOR APPLICATION NUMBER: 60/055, 984
;; PRIOR FILING DATE: 1997-08-18
;; PRIOR APPLICATION NUMBER: 60/058, 665
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: 60/058, 750
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: 60/058, 971
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: 60/058, 972
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: 60/058, 975
;; PRIOR FILING DATE: 1997-09-12
;; PRIOR APPLICATION NUMBER: 60/060, 834
;; PRIOR FILING DATE: 1997-10-02
;; PRIOR APPLICATION NUMBER: 60/060, 841
;; PRIOR FILING DATE: 1997-10-02
;; PRIOR APPLICATION NUMBER: 60/060, 844
;; PRIOR FILING DATE: 1997-10-02
;; PRIOR APPLICATION NUMBER: 60/060, 865
;; PRIOR FILING DATE: 1997-10-02
;; PRIOR APPLICATION NUMBER: 60/061, 059
;; PRIOR FILING DATE: 1997-10-02
;; PRIOR APPLICATION NUMBER: 60/061, 060
;; PRIOR FILING DATE: 1997-10-02
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 45
;; LENGTH: 1337
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (1335)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (1336)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (1337)
;; OTHER INFORMATION: n equals a,t,g, or c
;; US-09-209-462B-45

Query Match 72.1% Score 857; DB 16; Length 1337;
Best Local Similarity 99.8% Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCTCTCTGATTGCCCCGAACACCGTGACTGTTGTCGCCACAGAGG 199
|||||
Db 152 gccgcgcctctcttgatctgccccgaacacccgltgactctgctgcgcgacagagag 211
QY 200 CCGTGGTGTGAGAGCAATGGCCGATTCACCGGATCTGAGACCTGTTGAAACATCC 259
|||||
Db 212 cctgggtgtgtgagcgaaatgagcgaatctccacgcgaatccctggagccctgtgttaacatcc 271
QY 260 TCATCCCTGTGTAAACCGGATCGATATGTGAGAGTCTCAAGGAAATGTCTATCAACG 319
|||||
Db 272 tcaatccctgtgttagacaggaatccgaatagtcagagcttcaaggaatgtcatcaag 331
QY 320 TGCCTGAGCAGTGGCTGTGACTCTGACATCTACTGCAAAATCGATGAGTCTTT 379
|||||
Db 332 tgcctgagcagctgcgtgtgactctcgacaatgttaactctgcaaatcgaatgagctcttc 391
QY 380 ACCTGCGCATCATGAGACCTTACAAAGCAAGCTACGGTGTGAGAGACCTGACTATGCCG 439
|||||
Db 392 acctgcatcatagacccttacaagcaagctacggtgtgaggaacctgaglatgccc 451
QY 440 TCACCCAGCTAGCTCAAAACCATGATGATCAGAGCTCGGCAAACTCTNTGACAAAG 499
|||||
Db 452 tcaacccagctagctcaacaacacatgagatcagagctcgcaaacctctctgacaaag 511
QY 500 TCTTCCGGAACGAGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
|||||
Db 512 tcttcgaggaacggagctccctgaatgccaagctgtgagtcacatcaacaaagctgcgt 571
QY 560 ACTGCTGGGGTATCCCGCTGCTNCGTTATGAGATCAAGATATCATGTGCCACCCCGGG 619
|||||
Db 572 actgctgggtatccctgctctgctctatgagaatcaagatcatalgtgccaccccg 631
QY 620 TGAAGAGTATGACAGATCAGTGGAGGACAGGCGGCAAGGCGGCACAGTCTAG 679
|||||
Db 632 tgaagaagctcatgacagatgagagtcgaagtcgaagtcgaagtcgaagtcgaagtcgaag 691
QY 680 AGTCTGAGGGAGCCGAGAGTCCGCCATCATGTGGCAGAGGAGAAACAGGCCAGCA 739
|||||
Db 692 agctcagggagaccgagagctgagccatcaatgtgcaagagaggaagaaacagcccgag 751
QY 740 TCCCTGCGCTCCGAGCAAGAAAGCTGTAACAGATTAATCAGGACGAGAGCCAGCTG 799
|||||
Db 752 tccctgctcccgagcagaagaaagctgaacagaataatcagcagagagagagccagct 811
QY 800 CAGTTGCGGAGGAGGCAAGGCTAAAGCTGAAGCTATTTGCAATCTGGCTGCGAGCTCTGA 859
|||||
Db 812 cagttcggcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 871
QY 860 CACACATATATGAGATGACAGAGCTTCACTGACTGTGGCCGAGAGTATGTGACGCGCT 919
|||||
Db 872 cacaacataatgagatgcaagcagcttcaactgactgtgagcagcagatgtgcaagcagct 931
QY 920 TCTTCAAACTGGCCCAAGAGACTCCACACATATCTACTGCCCTCAAACTGGCGATGTCA 979
|||||
Db 932 tcttcaaacctgccaagagactccaacatacttacttgcctccaacccctgagagatca 991
QY 980 CCAGCATGCTGCTCAGGCGCATGCTATATGAGAGCCCTCACCAAGCCCGAGTGCAG 1039
|||||
Db 992 ccagcatgtgtgctcagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1051
QY 1040 GGACTCCAGACTCATCTCCAGTGGAGAGCAGAGATGTCCAGGGTACAGATGACAGT 1098
|||||
Db 1052 ggaactccagactcaactctccagctgagagcagcagagatgtccagggatcaagatgcaag 1110

RESULT 3
US-09-338-467-941
; Sequence 941, Application US/09338467
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

32 793 66.8 1416 22 US-09-598-075A-249 Sequence 249, App
33 728 61.3 1429 56 US-60-172-373-10483 Sequence 10483, A
34 674 56.7 1234 55 US-60-164-285-3468 Sequence 3468, App
35 582 49.0 732 33 US-09-878-134-349 Sequence 349, App
36 582 49.0 732 37 US-10-033-356-349 Sequence 349, App
37 572 48.1 1309 17 US-09-359-922-11863 Sequence 11863, A
38 572 48.1 1309 17 US-09-359-922-11863 Sequence 11863, A
39 509 42.8 622 36 US-09-998-598-1161 Sequence 1161, App
40 454 38.2 454 33 US-09-878-178-193 Sequence 193, App
41 454 38.2 454 37 US-10-046-935-193 Sequence 193, App
42 453 38.1 546 37 US-10-081-124-244 Sequence 244, App
43 453 38.1 550 17 PCT-US02-02870-2995 Sequence 2995, App
44 453 38.1 550 37 US-10-066-543-2995 Sequence 2995, App
45 417 35.1 483 25 US-09-652-127-341 Sequence 341, App

ALIGNMENTS

RESULT 1
PCT-US98-12125-45
Sequence 45, Application PC/TUS9812125

GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
NUMBER OF SEQUENCES: 318
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12125
FILING DATE: Jan 01, 1990
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P2008Complete
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US98-12125-45

Query Match 72.1% Score 857; DB 1; Length 1337;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCCCTCTCTGATGATGCCCCGAAACACCGTGTCTGTCGCGCCGACAGAGG 199
DB 152 GCCGCCCTCTCTGATGATGCCCCGAAACACCGTGTCTGTCGCGCCGACAGAGG 211
QY 200 CCTGGGTGTGAGCAATGGCGCATTCACCGGATCTCGAGACCTGGTTGAACATCC 259
DB 212 CCTGGGTGTGAGCAATGGCGCATTCACCGGATCTCGAGACCTGGTTGAACATCC 271

QY 260 TCATCCCTGTGTAGACCGGATCGATATGTGCAGAGTCCAGGAAATGTGATCAACG 319
DB 272 TCATCCCTGTGTGTAGACCGGATCGATATGTGCAGAGTCCAGGAAATGTGATCAACG 331
QY 320 TGCCTGACAGTGGCGCTGTGACTCTGCACAAATGTAATCTGCAAAATGATGAGTCTTT 379
DB 332 TGCCTGACAGTGGCGCTGTGACTCTGCACAAATGTAATCTGCAAAATGATGAGTCTTT 391
QY 380 ACCTGGCATCATGAGACCTTTACAGGCAAGCTACGGTGTGGAGAGACCTTGAGTCCG 439
DB 392 ACCTGGCATCATGAGACCTTTACAGGCAAGCTACGGTGTGGAGAGACCTTGAGTCCG 451
QY 440 TCACCCAGCTACGCTCAACCAACCATGATCAGAGTCCGGAACCTGTCNTTGACAAAG 499
DB 452 TCACCCAGCTACGCTCAACCAACCATGATCAGAGTCCGGAACCTGTCNTTGACAAAG 511
QY 500 TCCTCCGGGAACGGAGATCCCTGATGATCCAGCATTTGTGATGCCATCAACCACTGCTG 559
DB 512 TCCTCCGGGAACGGAGATCCCTGATGATCCAGCATTTGTGATGCCATCAACCACTGCTG 571
QY 560 ACTGCTGGGGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
DB 572 ACTGCTGGGGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 631
QY 620 TGAAGAAGTCTATGACATGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 679
DB 632 TGAAGAAGTCTATGACATGAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 691
QY 680 AGCTGAGAGGAGCCGAGAGTGGCCATCATGTGCGAGAGGAGGAGGAGGAGGAGGAGGAG 739
DB 692 AGCTGAGAGGAGCCGAGAGTGGCCATCATGTGCGAGAGGAGGAGGAGGAGGAGGAGGAG 751
QY 740 TCCTGGCTCGGAAGCAGAAAAAGGCTGAACAGATTAATCAGGACGAGAGGAGGAGGAG 799
DB 752 TCCTGGCTCGGAAGCAGAAAAAGGCTGAACAGATTAATCAGGAGCAGAGAGGAGGAGGAG 811
QY 800 CAGTTCTGGCGAAGCCCAAGCTTAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGA 859
DB 812 CAGTTCTGGCGAAGCCCAAGCTTAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGA 871
QY 860 CACAACATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 919
DB 872 CACAACATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 931
QY 920 TCTCCAACTGGCCAGAGACTCCAACTATCTACTGCTCCAACTGAGGATGCA 979
DB 932 TCTCCAACTGGCCAGAGACTCCAACTATCTACTGCTCCAACTGAGGATGCA 991
QY 980 CCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1039
DB 992 CCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
QY 1040 GGACTTCAGACTCACTCTCCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGT 1098
DB 1052 GGACTTCAGACTCACTCTCCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGT 1110

RESULT 2
US-09-209-462B-45
Sequence 45, Application US/09209462B

GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1
CURRENT APPLICATION NUMBER: US/09/209,462B
CURRENT FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 16:20:29 : Search time 2869.55 Seconds

(without alignments)
8956.896 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCCTTCGGAGCNAACCGCT.....GGAGCAGATTTCCTGATT 1188

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 21979536 segs, 10817449327 residues

Word size : 0

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```
Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US100_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
```

```
44: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	72.1	1337	1 PCT-US98-12125-45	Sequence 45, Appl
2	857	72.1	1337	16 US-09-209-462B-45	Sequence 45, Appl
3	857	72.1	1384	17 US-09-338-467-941	Sequence 941, App
4	857	72.1	1384	25 US-09-644-869-8024	Sequence 8024, App
5	857	72.1	1384	25 US-09-640-163-7309	Sequence 7309, App
6	857	72.1	1384	25 US-09-652-109-7925	Sequence 7925, App
7	857	72.1	1384	25 US-09-652-127-7358	Sequence 7358, App
8	857	72.1	1384	25 US-09-652-816-7419	Sequence 8809, App
9	857	72.1	1384	25 US-09-652-913-8809	Sequence 8809, App
10	857	72.1	1384	25 US-09-652-914-7997	Sequence 7997, App
11	857	72.1	1384	31 US-09-808-384-941	Sequence 941, App
12	857	72.1	1398	1 PCT-US01-03800A-1001	Sequence 1001, App
13	857	72.1	1398	1 PCT-US01-04058A-2212	Sequence 2212, App
14	857	72.1	1398	18 US-09-471-275-865	Sequence 865, App
15	857	72.1	1398	18 US-09-496-914A-8410	Sequence 8410, App
16	857	72.1	1398	22 US-09-560-875A-8410	Sequence 8410, App
17	857	72.1	1501	16 US-09-298-733-35	Sequence 35, Appl
18	857	72.1	1501	16 US-09-298-733A-35	Sequence 35, Appl
19	857	72.1	1501	29 US-09-723-594-35	Sequence 35, Appl
20	857	72.1	1501	29 US-09-724-497-35	Sequence 35, Appl
21	857	72.1	9098	1 PCT-US01-04942A-34	Sequence 34, Appl
22	857	72.1	9098	22 US-09-596-193A-37	Sequence 37, Appl
23	857	72.1	9278	1 PCT-US01-08656-5097	Sequence 5097, App
24	855	72.0	1322	1 PCT-US98-12125-102	Sequence 102, App
25	855	72.0	1322	16 US-09-209-462B-106	Sequence 106, App
26	794	66.8	1308	16 US-09-205-070-13379	Sequence 13379, App
27	794	66.8	1308	17 US-09-340-623-13379	Sequence 13379, A
28	794	66.8	1308	33 US-09-898-888-13379	Sequence 13379, A
29	794	66.8	1308	33 US-09-898-888A-13379	Sequence 13379, A
30	793	66.8	1416	1 PCT-US01-04038A-244	Sequence 244, App
31	793	66.8	1416	22 US-09-598-075-249	Sequence 249, App

Best Local Similarity	100.08;	Pred. No. 44;
Matches 17;	Conservative 0;	Mismatches 0;
Indels 0;	Gaps 0;	
Oy	554	CTGCTGACTGCTGGGT 570
Db	553	CTGCTGACTGCTGGGT 569

Search completed: September 22, 2002, 17:24:48
Job time: 7095 sec


```

Oy 1 GGCTTCTGGAGCANNACCGCTCCGCTCCTGCTTGGTTCCGGAGGTCGTGCGGGCGGTGG 60
Db 1 GGGTCTTCTGGAGCANNACCGCTCCGCTCCTGCTTGGTTCCGGAGGTCGTGCGGGCGGTGG 60
Oy 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
Oy 121 GGGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
Db 120 GGGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
Oy 180 GTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
Db 180 GTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
Oy 240 GGAGCCTGCTTGAACATCTCATCTCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCT 299
Db 240 GGAGCCTGCTTGAACATCTCATCTCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCT 299
Oy 300 CAAGGAATTTGTCATCAACGTCGCTGACAGCATCTGGCTGTGACTCTCGCAAAATGTACTCT 359
Db 300 CAAGGAATTTGTCATCAACGTCGCTGACAGCATCTGGCTGTGACTCTCGCAAAATGTACTCT 359
Oy 360 GCAATATCGATGGAGTCCTTCTTACCTGCGCATGTGAGACCCCTTACAAAGCAAGCTACGATGT 419
Db 360 GCAATATCGATGGAGTCCTTCTTACCTGCGCATGTGAGACCCCTTACAAAGCAAGCTACGATGT 419
Oy 420 GGAGGACCTGATGATGCGCGTACCCAGCTAGCTCAAAACAACATGAGATCAGAGCTCGG 479
Db 420 GGAGGACCTGATGATGCGCGTACCCAGCTAGCTCAAAACAACATGAGATCAGAGCTCGG 479
Oy 480 CAACTCTCTTGTGACAAAGCTTTCGGGAAAGGAGTCCCTGAATGCCAGATTTGTGGA 539
Db 480 CAACTCTCTTGTGACAAAGCTTTCGGGAAAGGAGTCCCTGAATGCCAGATTTGTGGA 539
Oy 540 TGCATCAACCAAGCTGCTGACTGCTGGGATATCCGCTGCTTATGATATCAAGA 599
Db 540 TGCATCAACCAAGCTGCTGACTGCTGGGATATCCGCTGCTTATGATATCAAGA 599
Oy 600 TATCATGCTGCCACCCCGGGGTGAAAAGTCTATGTGACAGATGCAAGGCGAGCGCGG 659
Db 600 TATCATGCTGCCACCCCGGGGTGAAAAGTCTATGTGACAGATGCAAGGCGAGCGCGG 659
Oy 660 GAAACGGGCGCAGTTCTAGAGTCTGAGGGAGCCCGAGAGTCCGATCAATGTGGCAGA 719
Db 660 GAAACGGGCGCAGTTCTAGAGTCTGAGGGAGCCCGAGAGTCCGATCAATGTGGCAGA 719
Oy 720 AGGAGAGAAACAGGCCCGCATCTGCTCGGAGCAGAGAAAGGCTGAACAGATAAATCA 779
Db 720 AGGAGAGAAACAGGCCCGCATCTGCTCGGAGCAGAGAAAGGCTGAACAGATAAATCA 779
Oy 780 GCGACGACGAGAGCGCAGTCTGCTGCGCAAGGCCAAGGCTAAAGCTAAGCTATTTCG 839
Db 780 GCGACGACGAGAGCGCAGTCTGCTGCGCAAGGCCAAGGCTAAGCTAAGCTATTTCG 839
Oy 840 AATCTGCTGCTCAGCTGTGACAAACATATGAGAGATGACAGCTTACTGACTGTGGC 899
Db 840 AATCTGCTGCTCAGCTGTGACAAACATATGAGAGATGACAGCTTACTGACTGTGGC 899
Oy 900 CGAGCAGTATGTCAAGCGCGCTTCTCAAACTGGCCAAAGAGCTCCAACTATCTACTGCG 959
Db 900 CGAGCAGTATGTCAAGCGCGCTTCTCAAACTGGCCAAAGAGCTCCAACTATCTACTGCG 959
Oy 960 CTCCAACTGCTGGCGATGTCCACAGCATGCTGAGCCATGTGGGTATGTGAGCCCT 1019
Db 960 CTCCAACTGCTGGCGATGTGTCCACAGCATGCTGAGCCATGTGGGTATGTGAGCCCT 1019
Oy 1020 CACCAAAAGCCCGAGTGCAGAGGAGCTCAGACTCTCCAGTGGGAGCAGCAGAGATGT 1079
Db 1020 CACCAAAAGCCCGAGTGCAGAGGAGCTCAGACTCTCCAGTGGGAGCAGCAGAGATGT 1079

```

```

Oy 1080 CCAGGCTACAGTACAGTNTTGATGAGAACTGTATCGAGTACAGATGATTAGTGAG 1139
Db 1080 CCAGGCTACAGTACAGTNTTGATGAGAACTGTATCGAGTACAGATGATTAGTGAG 1139
Oy 1140 CTGGGCTTNGCCAGGAGTCTGGGGACAAGAGCAGATTTCTGATTT 1188
Db 1140 CTGGGCTTNGCCAGGAGTCTGGGGACAAGAGCAGATTTCTGATTT 1188

```

RESULT 3

```

BC002442 1278 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, stomatin-like 2, clone MGC:1179 IMAGE:3346384, mRNA,
DEFINITION complete cds.
ACCESSION BC002442
VERSION BC002442.1 GI:12803254
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1278)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT: MGC help desk
 Email: cgaaps-remail.nih.gov
 Tissue procurement: ATCC/DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov

Shcherbako, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlinsky, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: m Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 9652258.

FEATURES

source

```

1..1278
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:1179 IMAGE:3346384"
/tissue_type="Skin, melanotic melanoma."
/clone_1ib="NTH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
47..1117
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH02442.1"
/db_xref="GI:12803255"
/translation="MLAARGTGALLRGLASGRAPRASSGLPNTVFLVPPO
EAWYRMGRFRILPEGLINTLIVLDIRIVOSIKELIVWPOSATVLDNATYLD
GVILRLINDPKASYGVDEPREYATVQLAOTMRSELGSLDKYFRESINATYDA
INQADMGIRCLREIKDHPVPRVSKMOMOVBAERKRAVLESGETRESINVA
EGKROAOILASEAKAEQINQAGEASVLAARAKAEVAIILAAALQHNQDAASL
TVAQVYSAFSKLAKDSNTILPSPNPDVTSMAQAMGVYALTKAPVPGTSDLSG

```

CDS

BASE COUNT 303 a 325 c 367 g 265 t
ORIGIN

Query Match 95.3%; Score 1132; DB 9; Length 1260;
Best Local Similarity 99.5%; Pred. No. 1.2e-258;
Matches 1154; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

30 TCGTGGTTCGAGAGTGTGCGGGGGGAAATGCGGGGGGGGGGGA 89
1 TCGTGGTTCGAGAGTGTGCGGGGGGAAATGCGGGGGGGGGGGA 59
90 CTGGGGCCCTTTTGTCTGAAGGGGCTCTACTGCTTGTGCGCGCTCGG -30CGCGCT 148
60 CTGGGGCCCTTTTGTCTGAAGGGGCTCTACTGCTTGTGCGCGCTCGGCGCGCT 119
149 CCTCTGAGATTTGCGCGGAAACACCGTGTACTTGTGCTGCGGAGAGAGGCTGGGTGG 208
120 CCTCTGAGATTTGCGCGGAAACACCGTGTACTTGTGCTGCGGAGAGAGGCTGGGTGG 179
209 TGGAGCGAATGGGGCGGATTTCCACCGGATCTGGAGCTGGTTTGAACATCTCATCCCTG 268
180 TGGAGCGAATGGGGCGGATTTCCACCGGATCTGGAGCTGGTTTGAACATCTCATCCCTG 239
269 TGTAGACCGGATTCGAGATGTGACAGAGTCTCAGAGAAATTTGTCATCAACCTGCGTGAAGC 328
240 TGTAGACCGGATTCGAGATGTGACAGAGTCTCAGAGAAATTTGTCATCAACCTGCGTGAAGC 299
329 AGTGGGCTGTACTCTGCAATGTAATCTTGCAGAAATGATGAGTCTTTTACTTGGCA 388
300 AGTGGGCTGTACTCTGCAATGTAATCTTGCAGAAATGATGAGTCTTTTACTTGGCA 359
389 TCATGACCGCTTCAAGAGCAAGCTGAGGTGTGAGAGACCGTGTATGCGCGTCAACCGCA 448
360 TCATGACCGCTTCAAGAGCAAGCTGAGGTGTGAGAGACCGTGTATGCGCGTCAACCGCA 419
449 TAGCTCAACAACCATGTGAGATCAGAGCTCGGCAAACTCTCTTGTGACAAGTCTTCCGGG 508
420 TAGCTCAACAACCATGTGAGATCAGAGCTCGGCAAACTCTCTTGTGACAAGTCTTCCGGG 479
509 AACGGAGTCCCTGAAATGCGGATTTGGATGCGCATCAACCAAGCTGTGTACTGTGGG 568
480 AACGGAGTCCCTGAAATGCGGATTTGGATGCGCATCAACCAAGCTGTGTACTGTGGG 539
569 GTATCGGCTGCTTNGTATGATGATCAAGATATCATGTGCGCACCGCGGTGAAAGAT 628
540 GTATCGGCTGCTTNGTATGATGATCAAGATATCATGTGCGCACCGCGGTGAAAGAT 599
629 CTATGACATGCAAGTGTGAGGAGGCGGCGGAAACAGTCTTGAAGTCTGAGG 688
600 CTATGACATGCAAGTGTGAGGAGGCGGCGGAAACAGTCTTGAAGTCTGAGG 659
689 GGACCCGAGATCGGCGCATCATGTGCGGAGAGGAGGAGGCGGAGTCTGCGCT 748
660 GGACCCGAGATCGGCGCATCATGTGCGGAGAGGAGGAGGCGGAGTCTGCGCT 719
749 CGAAGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTCTGCGCT 808
720 CGAAGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTCTGCGCT 779
809 CGAAGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTCTGCGCT 868
780 CGAAGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTCTGCGCT 839
869 ATGAGGATGCAAGCTGCTGAGTGTGCGGAGGAGGAGTGTGAGGCGGCTTCCAAAC 928
840 ATGAGGATGCAAGCTGCTGAGTGTGCGGAGGAGGAGTGTGAGGCGGCTTCCAAAC 899
929 TGGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTGTGAGGCGGCTTCCAAAC 988
900 TGGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTGTGAGGCGGCTTCCAAAC 959
989 TGGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTGTGAGGCGGCTTCCAAAC 1048

Db 960 TGGCGAAGAGGCTGAAGATTAATCAGGAGGAGGAGGCGGAGTGTGAGGCGGCTTCCAAAC 1019
Qy 1049 ACTCACTCTCCAGTGGGAGGAGGAGGAGTGTCCAGGATGCAATGCAAGTNTTGTAGAG 1108
Db 1020 ACTCACTCTCCAGTGGGAGGAGGAGGAGTGTCCAGGATGCAATGCAAGTNTTGTAGAG 1079
Qy 1109 AACTTGATGAGTCAAGATGAGTGTGAGGCTGGGCTTNGCAGGAGGAGTCTGGGGAGCA 1168
Db 1080 AACTTGATGAGTCAAGATGAGTGTGAGGCTGGGCTTNGCAGGAGGAGTCTGGGGAGCA 1139
Qy 1169 GGAAGCAGATTTTCTGATT 1188
Db 1140 GGAAGCAGATTTTCTGATT 1159
RESULT 5
AK027405 1244 bp mRNA linear PRI 15-MAY-2001
LOCUS Homo sapiens CDNA FLJ14499 fis, clone NT2RM1000080, weakly similar
DEFINITION To UNC-1 PROTEIN.
ACCESSION AK027405
VERSION AK027405.1 GI:14042059
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 CDNA to mRNA, clone.lib:NT2RM1 clone:NT2RM1000080.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Magatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murekami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakematsu,A., Nakamura,Y., Nagahara,K., Masuko,Y., Ninomiya,K. and Iwayanagi,T. NEBO human cDNA sequencing project Unpublished
2 (bases 1 to 1244)
Isogai,T. and Otsuki,T.
REFERENCE 10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp. Tel:81-438-52-3951. Fax:81-438-52-3952)
COMMENT NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA library construction; Research Association for Biotechnology; cDNA library construction, 5' and 3' end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
source
Location/Qualifiers
1..1244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM1000080"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM1"
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells."
35..1105
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55091.1"
/db_xref="GI:14042060"
/translation="MLARARGTGALLRGLASGRAPRRASSGLPNTVLPVPOO EAWVERGRHRHRIIEPGLNIIIPVDIRIVOSKEIVINWPDSATLTNDYLOID GVLIRIMDPYKASTGVEDPEYATVOLAQTTRSLGLSLDKVFERESLNASIVDA INQADWGIKRLREYEDIDHVPVPRVSKMGOVEAERKRATVLESGTRESAIVNA EGKROAOILASEARKEAOINOAGSASVILKAKAKAIAIILAAALPHOHNGDAASL TVAEQVSAFSKLAKDSNTILLPSNPGDVTSMVAQAMGVGALTRKAPVPGPDSLSGS

BASE COUNT 284 a 327 c 368 g 265 t
ORIGIN

Query Match 95.28; Score 1130.4; DB 9; Length 1244;
Best Local Similarity 99.48; Pred. No. 2, 9e-258;
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

30 TCGTGGTTCCGGAGCGTGGCGGGCGGTGGAATCTGGCGCGCGCGGAGGCA 89
1 TCGTGGTTCCGGAGCGTGGCGGGCGGTGGAATCTGGCGCGCGCGGAGGCA 59
90 CTGGGGCCCTTTTCTGCTGAGGGGCTCTACTGCGCTTGGCGCGCGCTCG- 148
60 CTGGGGCCCTTTTCTGCTGAGGGGCTCTACTGCGCTTGGCGCGCGCGCTCG 119
149 CCTCTGGATTGCCCCGCAACACCGTGTACTGTTCTGTCGGCAGCAGAGGCGCTGG 208
120 CCTCTGGATTGCCCCGCAACACCGTGTACTGTTCTGTCGGCAGCAGAGGCGCTGG 179
209 TGGAGCGAATGGGGCGATTCCACCGGATCCGAGCGCTGGTTTGAACATCCTCATCCCTG 268
180 TGGAGCGAATGGGGCGATTCCACCGGATCCGAGCGCTGGTTTGAACATCCTCATCCCTG 239
269 TGTTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACGTCCTGAGC 328
240 TGTTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACGTCCTGAGC 299
329 AGTGGCGTGTGACTCTGAGCAATTAATCTGCAAAATGAGAGGAGTCTTACTGGGCA 388
300 AGTGGCGTGTGACTCTGAGCAATTAATCTGCAAAATGAGAGGAGTCTTACTGGGCA 359
389 TCATGAGACCCCTTACAGAGCAAGTACGCTGTGAGAGACCCCTGAGTATGCCCTCAACCC 448
360 TCATGAGACCCCTTACAGAGCAAGTACGCTGTGAGAGACCCCTGAGTATGCCCTCAACCC 419
449 TAGCTCAAAACACCATGATGATCAAGTCCGCAAACTCTCTNTGACAAATCTTCCGGG 508
420 TAGCTCAAAACACCATGATGATCAAGTCCGCAAACTCTCTNTGACAAATCTTCCGGG 479
509 AAGGGAGTCCCTGATATGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 568
480 AAGGGAGTCCCTGATATGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 539
569 GTATCCGCTGCTTCGCTTATGATGATCAAGATATCATGTCACACCCCGGGTGAAGAGT 628
540 GTATCCGCTGCTTCGCTTATGATGATCAAGATATCATGTCACACCCCGGGTGAAGAGT 599
629 CTATGACAGATGACAGTGTGAGAGCAAGCGCGGAAACCGGCGCAGATTCTAGAGTGTGAG 688
600 CTATGACAGATGACAGTGTGAGAGCAAGCGCGGAAACCGGCGCAGATTCTAGAGTGTGAG 659
689 GGACCCGAGATGCGGCGATCATATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748
660 GGACCCGAGATGCGGCGATCATATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
749 CCGAAGCAGAAAAGGCTGACAGATTAATCAAGCAGCAGAGAGCGCAGTGTGAGT 808
720 CCGAAGCAGAAAAGGCTGACAGATTAATCAAGCAGCAGAGAGCGCAGTGTGAGT 779
809 CGAAGGCCAAGGCTAAAGCTGAAGCTATTCGAAATTCCTGGCTGACCTGCAACAACTA 868
780 CGAAGGCCAAGGCTAAAGCTGAAGCTATTCGAAATTCCTGGCTGACCTGCAACAACTA 839
869 ATGAGATGACAGACAGTCTGATGATGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
840 ATGAGATGACAGACAGTCTGATGATGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 899
929 TGGCCAAAGAGACTCAACACTATCTACTGCTTCCACCTGGCGAGTGTACACAGCATG 988
900 TGGCCAAAGAGACTCAACACTATCTACTGCTTCCACCTGGCGAGTGTACACAGCATG 959

QY 989 TGGCTCAGGCCATGGGTGTATATGAGACCCCTCAACAAAGCCCGAGGCGAGCTCAG 1048
DB 960 TGGCTCAGGCCATGGGTGTATATGAGACCCCTCAACAAAGCCCGAGGAGCTCAG 1019
QY 1049 ACTCAGCTCCAGTGGAGAGCAGAGATGTCCAGGGGTACAGATGCAAGTNTGTAGAG 1108
DB 1020 ACTCAGCTCCAGTGGAGAGCAGAGATGTCCAGGGGTACAGATGCAAGTNTGTAGAG 1079
QY 1109 AACTGTAGTACAGTCAAGATGATGATGAGAGCTGGGCTTGGCAGGAGAGTGTGGGCA 1168
DB 1080 AACTGTAGTACAGTCAAGATGATGATGAGAGCTGGGCTTGGCAGGAGAGTGTGGGCA 1139
QY 1169 GGAAGCAGATTTTCTGATT 1188
DB 1140 GGAAGCAGATTTTCTGATT 1159

RESULT 6

BC010152 1281 bp mRNA 11near PRI 12-JUL-2001
LOCUS
DEFINITION
complete cds.
ACCESSION
BC010152
VERSION
BC010152.1 GI:14603402
KEYWORDS
MGC.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1281)

REFERENCE

AUTHORS
Strausberg, R.
JOURNAL
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rudin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McKloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J.,
Tjongson, E.E., Touchman, J.W., Tsurgoev, C., Vogt, D.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Series: IRAL Plate: 27 Row: d Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

source

1..1281
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:19715 IMAGE:3534656"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NHL MGC-7"
/lab_host="DH10B-R"
/note="Vector: pOT87"
44..1114
/codon_start=1

CDS

```

/product="stomatin-like 2"
/protein_id="AAH10152.1"
/db_xref="GI:14603403"
/translation="MLAARAGCTGALLIRGLISAGRAPRHSSGLPNTVYLFVPOO
EAMVYERMGRRHRLERBLNLIIVYDRIIRYVLSKEIVYIPDQSAITDNTLQID
GVLTLRIMDPYKASYGVEDPEYAVTQPPQATVMSRELGLSDKFRRESLNASIVDA
INQADWGIJNCLREYKIDIHPRVKESSMOMQVAEIRKATVLESGTRESINVA
EGKKQAOIILASEAKAEQIINOAGEASVILKAKAKAEKAIILAAALTOHNGDAASL
TVAEQYSAFSLKAKDSNTLILPSNPQVTSMAVQAGVYALIKAPVPGTSDLSGS
SSRDVQGTDSLDELDRVKNMS"
BASE COUNT      315 a      330 c      372 g      264 t
ORIGIN

```

```

Query Match      95.2% Score 1130.4; DB 9; Length 1281;
Best Local Similarity 99.4%; Pred. No. 2.9e-258;
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```

```

QY 30 TCGTTGGTTCCGGAGTGCCTGCGCGGTGGGAAATGCTGGCGCGCGCGGAGGCA 89
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 TCGTTGGTTCCGGAGTGCCTGCGCGGTGGGAAATGCTGGCGCGCGCGGAGGCA 68
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 90 CTGGGGCCCTTTTGTGAGGGGCTCTACTGCTTGTGGCGGCTCCG-CCGCGCCT 148
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 CTGGGGCCCTTTTGTGAGGGGCTCTACTGCTTGTGGCGGCTCCGCGCGCCT 128
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 149 CCTGTGATGCCCCGAAACCCGCTGACTGTGCTGCCGACAGAGAGGCTGGGTGG 208
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 CCTGTGATGCCCCGAAACCCGCTGACTGTGCTGCCGACAGAGAGGCTGGGTGG 188
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 209 TGAACGGAATGGGCGGATTTCCACCGGATCTGGAGCTGTGTTGAACATCCCTCATCCCTG 268
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 TGAACGGAATGGGCGGATTTCCACCGGATCTGGAGCTGTGTTGAACATCCCTCATCCCTG 248
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 TGTTAGACCGGATCCGATATGTGACAGATCTCAAGAAATTTCTCAACAGTGGCTGAC 328
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 249 TGTTAGACCGGATCCGATATGTGACAGATCTCAAGAAATTTCTCAACAGTGGCTGAC 308
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 329 AGTGGGCTGTGACTCTGCACATGTAACCTGCAATGAGATGAGATCCCTTACCTGGCGCA 388
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 AGTGGGCTGTGACTCTGCACATGTAACCTGCAATGAGATGAGATCCCTTACCTGGCGCA 368
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 389 TCATGAGACCTTTAAAGGCAAGCTACGGTGTGGAGAGACCTGAGTATGCCCTCACCCAGC 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 369 TCATGAGACCTTTAAAGGCAAGCTACGGTGTGGAGAGACCTGAGTATGCCCTCACCCAGC 428
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 449 TAGCTCAAAACACATGAGATCAGAGCTCGCAAACTCTCTTTGGAACAAAGTCTTCCGG 508
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 429 CAGCTCAAAACACATGAGATCAGAGCTCGCAAACTCTCTTTGGAACAAAGTCTTCCGG 488
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 509 AACGGGATCCCTGAAATCCAGCATTTGGAGTGCATCAACCAAGCTGTAAGTCTGCTGGG 568
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 489 AACGGGATCCCTGAAATCCAGCATTTGGAGTGCATCAACCAAGCTGTAAGTCTGCTGGG 548
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 569 GTATCCGCTGCTTNCCTTATGATCAAGATATCATGTGCCACCCGCGTGAAGAGAT 628
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 549 GTATCCGCTGCTTNCCTTATGATCAAGATATCATGTGCCACCCGCGTGAAGAGAT 608
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 629 CTATGCAATGAGTGTGAGAGCAGAGCGCGGAAACGGGCCACAGTTCTAGAGTCTGAGG 688
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 609 CTATGCAATGAGTGTGAGAGCAGAGCGCGGAAACGGGCCACAGTTCTAGAGTCTGAGG 668
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 689 GGACCCGAGAGTGGCCATCATGTGGCAGAAAGGGAAGAACAGGCCAGATCCGGGCT 748
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 669 GGACCCGAGAGTGGCCATCATGTGGCAGAAAGGGAAGAACAGGCCAGATCCGGGCT 728
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 749 CCGAAGCAGAAAAAGCTAGACAGATTAATCAAGCAGACAGAGAGGCCAGTGCAGTTCTGG 808
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 729 CCGAAGCAGAAAAAGCTAGACAGATTAATCAAGCAGACAGAGAGGCCAGTGCAGTTCTGG 788
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 809 CGAAGGCCAAGGCTAAAGCTGAAGCTATTTCAGATCTCTGGCTCAGCTCTGACACAACATA 868
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 789 CGAAGGCCAAGGCTAAAGCTGAAGCTATTTCAGATCTCTGGCTCAGCTCTGACACAACATA 848
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 869 ATGAGAGTCCACAGCTTTCAGTACGTGGCCGAGCAGATGTGACAGCGGCTTCCAAAC 928
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 849 ATGAGAGTCCACAGCTTTCAGTACGTGGCCGAGCAGATGTGACAGCGGCTTCCAAAC 908
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 929 TGGCCAAAGAGCTCAACACTATATCTACTGCTCCACCAACCTTGGCATGTCAACAGCATGG 988
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 909 TGGCCAAAGAGCTCAACACTATATCTACTGCTCCACCAACCTTGGCATGTCAACAGCATGG 968
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 989 TGGCTCAGAGCCATGGGTGTATGTGAGCCCTTACCAGAGCCCAATGTCAGGAGTCCAG 1048
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 969 TGGCTCAGAGCCATGGGTGTATGTGAGCCCTTACCAGAGCCCAATGTCAGGAGTCCAG 1028
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1049 ACTCACTCCAGTGGAGCAGACAGATGTCCAGAGTACAGATCAAGTNTTGAATGAG 1108
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1029 ACTCACTCCAGTGGAGCAGACAGATGTCCAGAGTACAGATCAAGTNTTGAATGAG 1088
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1109 AACTGTATGAGTCAAGATGATGAGTGGGCTTGGCCAGGAGTCTGGGGACAA 1168
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1089 AACTGTATGAGTCAAGATGATGAGTGGGCTTGGCCAGGAGTCTGGGGACAA 1148
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1169 GGAACAGATTTTCTGATT 1188
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1149 GGAACAGATTTTCTGATT 1168
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 7

```

BC014990
LOCUS      BC014990      1262 bp      mRNA      linear      PRI 29-OCT-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:23225 IMAGE:4907745, mRNA, complete cds.
ACCESSION BC014990
VERSION    BC014990.1 GI:15929069
KEYWORDS   MGC.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1262)
AUTHORS   Strausberg, R.
TITLE      Direct Submission
JOURNAL    Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

```

```

REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

```

```

Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kuesche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Ilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalins, Michael Smith, Loraine Spence, Jeff Stolt,
Michael Thorne, Miranada Teal, Nastasia van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3-E. Consortium/LLNL at: http://image.llnl.gov
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

```

```

FEATURES
location/Qualifiers
1..1262
/organism="Homo sapiens"
/db_xref="LocusID:30968"

```

```

/db_xref="taxon:9606"
/clone="MGC:21225 IMAGE:4907745"
/tissue="type="Pancreas", epithelioid carcinoma"
/clone_lib="NIH_MGC_42"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
33..1103
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH14990.1"
/db_xref="GI:15929070"
/translation="MLAKARRGTCALLRLKSLASGRAPRRASGLPRNTVYLFVPOQ
EAWVERMGRFHRILPEPLNLILPVDLRIRVOSLKIIVNPQSAVTLNVTLOID
GVYLRIMDPYKASYGVEDEYAVTOLAQTMRSELKRLDKYFRERESINAVIA
INQADMGIRCLRYEIKDIHPRVRESQMOYEAERKRPATVLESEGTRESAINVA
EKGKQAOILASEAKAEQIINOAGESAVALAKAKAEATRIILAAVLOHNGDAASL
TVAEQVSAFSKLAKDSNTLILPSPNDVYSMAQAMGVGALIKAPVPGIPDLSLSSG
SSRDVQCTDASLDELDRLVKNKS"
BASE COUNT      305 a      326 c      367 g      264 t
ORIGIN
Query Match      95.1% Score 1130: DB 9: Length 1262:
Best Local Similarity 99.5% Pred. No. 3.6e-258;
Matches 1152: Conservative 0: Mismatches 4: Indels 2: Gaps 2:
QY 32 GTTGATTCGGAGGTCGTGGCGGCGGTGGCAAAATGCTGCGCGCGCGCGGCGGCACT 91
Db 1 GTTGATTCGGAGGTCGTGGCGGCGGTGGCAAAATGCTGCGCGCGCGCGGCGGCACT 59
QY 92 GGGGCGCTTTTGTCTGAGGGGCTCTACTGGCTTGTGGCGCGGCTTCG-GCGCGCGCTTC 150
Db 60 GGGGCGCTTTTGTCTGAGGGGCTCTACTGGCTTGTGGCGCGGCTTCG-GCGCGCGCTTC 119
QY 151 TCTGGATTCGCCGGAACACCGGTGTAATGTTCTGTCGCCGACAGAGAGCGCTGGGTGTG 210
Db 120 TCTGGATTCGCCGGAACACCGGTGTAATGTTCTGTCGCCGACAGAGAGCGCTGGGTGTG 179
QY 211 GAGGGAATGGCGCGATTCACCGGATCCTGGAGCGCTGGTTGAACATCCTATCCCTGTG 270
Db 180 GAGGGAATGGCGCGATTCACCGGATCCTGGAGCGCTGGTTGAACATCCTATCCCTGTG 239
QY 271 TTAAACCGGATTCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACAGTGGCTAGAGC 330
Db 240 TTAAACCGGATTCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACAGTGGCTAGAGC 299
QY 331 TCGGCTGTGACTTCGACAAATGTAATCTGCAAAATCGATGAGTCTTACCTGGCATC 390
Db 300 TCGGCTGTGACTTCGACAAATGTAATCTGCAAAATCGATGAGTCTTACCTGGCATC 359
QY 391 ATGACACCTTACAAAGGCAAGGTACGGTGTGAGAGACCTGATGATGCGGTACCCAGCTA 450
Db 360 ATGACACCTTACAAAGGCAAGGTACGGTGTGAGAGACCTGATGATGCGGTACCCAGCTA 419
QY 451 GCTAAACCAACCATGAGATTCAGAGTCTGGCAAACTCTTNTGGAACAAAGTCTTCCGGGAA 510
Db 420 GCTAAACCAACCATGAGATTCAGAGTCTGGCAAACTCTTNTGGAACAAAGTCTTCCGGGAA 479
QY 511 CGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGTGTGATGCTGGGGT 570
Db 480 CGGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGTGTGATGCTGGGGT 539
QY 571 ATCCGCTCCCTNCTTATGAGATCAAGATATTCATGTGCCACCCGGGTGAAGAGTCT 630
Db 540 ATCCGCTCCCTNCTTATGAGATCAAGATATTCATGTGCCACCCGGGTGAAGAGTCT 599
QY 631 ATGCAATGCAAGTGTGAGAGGCGGCGGGAACCGGCGCACAGTTCTAGAGTCTAGGGG 690
Db 600 ATGCAATGCAAGTGTGAGAGGCGGCGGGAACCGGCGCACAGTTCTAGAGTCTAGGGG 659
QY 691 ACCGAGAGTGGCGCATCAATGTGGCAGAGGGAAGAAACAGGCCAGATCCTGGCCCTCC 750
Db 660 ACCGAGAGTGGCGCATCAATGTGGCAGAGGGAAGAAACAGGCCAGATCCTGGCCCTCC 719

```

```

QY 751 GAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGGCCAGTGTCTGGCG 810
Db 720 GAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGGCCAGTGTCTGGCG 779
QY 811 AAGGCCAAGGCTAAAGCTGAAGCTATTCTGATCTCTGCTGCAGCTCTGACACAAATAT 870
Db 780 AAGGCCAAGGCTAAAGCTGAAGCTATTCTGATCTCTGCTGCAGCTCTGACACAAATAT 839
QY 871 GGAGATGAGAGAGCTTCACTGACTGTGGCGGAGAGTATGTACAGGCGCTTCCAAACTG 930
Db 840 GGAGATGAGAGAGCTTCACTGACTGTGGCGGAGAGTATGTACAGGCGCTTCCAAACTG 899
QY 931 GCCAAGAGCTCCACACTATCTACTGCGCTCCAAACCTGGCGATGTCACAGCATGTG 990
Db 900 GCCAAGAGCTCCACACTATCTACTGCGCTCCAAACCTGGCGATGTCACAGCATGTG 959
QY 991 GCTCAGGCGATGGTGTATATGAGAGCCTCAACCAAGCCCAAGTCCAGGACATCCAGAC 1050
Db 960 GCTCAGGCGATGGTGTATATGAGAGCCTCAACCAAGCCCAAGTCCAGGACATCCAGAC 1019
QY 1051 TCACCTCTCCAGTGGGAGGAGCAGAGATGTCAGAGGTACAGATGCAAGTNTGATGAGGAA 1110
Db 1020 TCACCTCTCCAGTGGGAGGAGCAGAGATGTCAGAGGTACAGATGCAAGTNTGATGAGGAA 1079
QY 1111 CTGTATCAGTCAAGATGAGTTAGTGAGAGCTGGGCTTNGCCAGGAGAGTCTGGGCAAGG 1170
Db 1080 CTGTATCAGTCAAGATGAGTTAGTGAGAGCTGGGCTTNGCCAGGAGAGTCTGGGCAAGG 1139
QY 1171 AAGCAGATTTTCTGATAT 1188
Db 1140 AAGCAGATTTTCTGATAT 1157

```

```

RESULT 8
BC003025
LOCUS      1257 bp      mRNA      linear      PRI 12-JUL-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:4191 IMAGE:2821269, mRNA,
complete cds.
ACCESSION      BC003025.
VERSION      BC003025.1 GI:12804332
KEYWORDS      MGC.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1257)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (06-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCRD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandon, Anna-Liisa Prabhui, Parvaneh Saeedi, Jacqueline
Schelin, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stolt,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>


```

/translaction="MLARAARGTALLRSGVQASGRVPRRASSGLPNTVILFVPOO
EAMVERMKRGRHRIIEBGLNVLIPDLIRIRVOSLKEIVINVPQASVTLNDVLIQID
GLVLRIMDPYKASYGVEDAVTQALRSELSLSDKFRRESESNANIVDA
INQADCMQJRCLEIKDIDHPPRVKESOMOEKBRKATVLESGTESAIVNA
EGKMOAOILASEAKAEIOINQAGEASAVLAKAKAEKRIIILAGATLONHGDAAASL
TVAKOIVSARSKLKLDKSNVTLPSNPDSVTSMAVQANGVIALTKAPVPGKONSQSR
RDVQATDTSLEELRKLIS"
BASE COUNT      320 a      311 c      364 g      258 t
ORIGIN

```

```

Query Match      73.0%: Score 867.4; DB 10; Length 1253;
Best Local Similarity 88.0%: Pred. No. 9.7e-196;
Matches 1002; Conservative 0; Mismatches 125; Indels 12; Gaps 5;

QY 34 TGGTTCGGAGGCTGCTGCGGGTGGGAAATGCTGCGCGCGCGCGGGGAGCTGG 93
DB 2 TGGTTCGGAGGCTGCGCGGGTGGGAAATGCTGCGCGCGCGGGGAGCTGG 60

QY 94 GGGCTTTTGTCTGAGAGGCTCTCTACTGCTTGGCCGCGCTCGG-6CCGCGCTCTC 152
DB 61 AGCGCTTGTCTGAGAGGCTCTCTACTGCTTGGCCGCGCTCGGCGCGCGCTCTC 120

QY 153 TGGATTTGCCCGAAACACCGTGTACTGCTGCTGCGCGAGAGAGCGCTGGGTGGTGA 212
DB 121 TGGATTTGCCCGAAACACCGTGTACTGCTGCTGCGCGAGAGAGCGCTGGGTGGTGA 180

QY 213 GCGAATGGGCGGATTCACCGGATCTGAGAGCTGTTGAACATCCATCTGCTGT 272
DB 181 GCGAATGGGCGGATTCACCGGATCTGAGAGCTGTTGAACATCCATCTGCTGT 240

QY 273 AGACCGGATCCGATATGTGACAGAGTCTCAAGAAATTTGTCATCAACGTGCTGAGCAGTC 332
DB 241 AGACCGGATCCGATATGTGACAGAGTCTCAAGAAATTTGTCATCAACGTGCTGAGCAGTC 300

QY 333 GGGCTGTGCTCTCGCATGTAACTCTGCAATTCATGAGAGTCTTACCTGCGCATCAT 392
DB 301 AGCGCTTGTCTGAGAGGCTCTGCAATTCATGAGAGTCTTACCTGCGCATCAT 360

QY 393 GGACCTTTTACAGGCAAGCTACGCTGTGAGAGAGCCCTGATGTCGCTGACCCAGCTAGC 452
DB 361 GGATTCCTTACAGGCAAGCTACGCTGTGAGAGAGCCCTGATGTCGCTGACCCAGCTAGC 420

QY 453 TCACCAACACATGAGATGAGAGCTGCGCAAACTCTCTTGACAAAGTCTTCCGGAGAGC 512
DB 421 TCAGAGCTGCTGAGATGAGAGCTGCGCAAACTCTCTTGACAAAGTCTTCCGGAGAGC 480

QY 513 GGAGTCCCTGATGCGCAGCATTTGTGATGCCATCAACCAAGTCTGCTGCTGGGCTAT 572
DB 481 TGAGTCCCTGATGCGCAGCATTTGTGATGCCATCAACCAAGTCTGCTGCTGGGCTAT 540

QY 573 CCGGTGCTTGTGATGATCAAGATATCCATGTGCCACCCCGGATGAAGAAGTCTAT 632
DB 541 CCGATGCTTGTGATGATCAAGATATCCATGTGCCACCCCGGATGAAGAAGTCTAT 600

QY 633 GCAGATGCGAGTGAAGGAGAGCGCGGAAACGCGCCACAGTTCTAGAGTCTGAGGGAGC 692
DB 601 GCAGATGCGAGTGAAGGAGAGCGCGGAAACGCGCCACAGTTCTAGAGTCTGAGGGAGC 660

QY 693 CCGAGTGGCGCATCATGTGTGGCAGAGAGGAAAGAAACAGCGCCAGATCTGGCTCCGA 752
DB 661 AGCAGAGTGGCGCATCATGTGTGGCAGAGAGGAAAGAAACAGCGCCAGATCTGGCTCCGA 720

QY 753 AGCAAAAAGGCTGAACAGATTAATCAGGAGCAGAGAGCGCTGAGTTCTGCGGAA 812
DB 721 AGCAAAAAGGCTGAACAGATTAATCAGGAGCAGAGAGCGCTGAGTTCTGCGGAA 780

QY 813 GGGCAGGCTAAAGCTGAGATCTATTCGATCTGCTGCTGAGCAACAACATTAATGG 872
DB 781 GGGCAGGCTAAAGCTGAGATCTATTCGATCTGCTGCTGAGCAACAACATTAATGG 840

QY 873 AGATGACAGAGTTTACGAGCTGTGGCCGAGAGATATGTCAAGCGCTTCTCAAACTGGC 932
DB 840 AGATGACAGAGTTTACGAGCTGTGGCCGAGAGATATGTCAAGCGCTTCTCAAACTGGC 900

```

```

DB 841 AGATGACAGAGTTTACGAGCTGTGGCCGAGAGATATGTCAAGCGCTTCTCAAACTGGC 900
QY 933 CAAGAGCTCAACACTATTCCTACTGCTCCCTCCCAACCTGGAGATGTCACAGCAGTGGGC 992
DB 901 CAAGATTTCCACACAGAGTCTACTGCTCCCTCCCAATCCAGCGAGCTCAGAGATGATGGC 960
QY 993 TCAGGCTAGGCTGTATATGAGAGCCCTCAACCAAGCCCAAGTGGCAGGAGTCTGAGCTC 1052
DB 961 TCAGGCTAGGCTGTATATGAGAGCCCTCAACCAAGCCCAAGTGGCAGGAGTCTGAGCTC 1014
QY 1053 ACTTCCAGTGGGAGCAGCAGAGATGTCAGAGTTCACAGATGCAATGATGAGAGT 1112
DB 1015 GAATCTCAGCGCAGAGAGAGATGTCAGAGTTCACAGAGTTCAGAGAGTAT--TGAAGAACT 1071
QY 1113 TGAATGAGTCAAGATGATGATGAGTGGAGCTTGGCAGGAGTCTGGGAGCAAGA 1171
DB 1072 GGGTATGAGTCAAGTCAAGTATGATGAGAT--GACTGCGCAGGAGATCTAGGGCGCGGA 1129

RESULT 10
BC003425
LOCUS
DEFINITION
MUS musculus, RIKEN cDNA 0610038F01 gene, clone MGC:6777
IMAGE:2609584, mRNA, complete cds.
ACCESSION
BC003425
VERSION
BC003425.1 GI:13097353
KEYWORDS
MGC.
SOURCE
house mouse.
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Strausberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNI)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: Villalobos, tmc.edu.
Villalobos, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/HLNI at: http://image.llnl.gov
Series: IRAC Plate: 5 Row: K Column: 5.
Location/Qualifiers
1. 1565
/organism="Mus musculus"
/db_xref="LocusID:66592"
/db_xref="Taxon:10090"
/clone="MGC:6777 IMAGE:2609584"
/tissue-type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/Note="Vector: pCMV-SPORT6"
12. 1073
/codon_start=1
/product="RIKEN cDNA 0610038F01 gene"
/protein_id="AAH03425.1"
/db_xref="GI:13097353"
/translaction="MLARAARGTALLRSGVQASGRVPRRASSGLPNTVILFVPOO

```

CDs

OY	619	GTGAAG-AGTCATGCAAGATGCAAGTGGAGGACAGAC-GGGCGAAACGGGCGACAGTTC	676
Db	421	GTGAAGAAGCTTAGCANNATGCAAGTGGAGGCAAGGGGGGAAACGGGGCAAGTTC	480
OY	677	TAGAGTCTGAGGGAGCCGAGAGTGGCCATCAATGTGGCAGAAAG-----GAAAGAAACA	731
Db	481	TAGAGTCTGAGGGAGCCGGAAGTGGGNCATCAATGTGGCGAANNAAGGAAGAAAGCAAG	540
OY	732	GGCCAGATCCTGGGCT--CCGAGACAGAAAAGCTGAACAGA-TAAATCAGGCAAGCAGG	788
Db	541	GGCCAGATCCTGGGCTTCCNNAACAAAAAAGGNTGAACAAANTPAATCAGGCACCGAG	600
OY	789	AG-AGGCCAGTGCATTTGCGCGAAGGCCAAGGCTAAAGCTGAAGCTTATCGAATCTGG	847
Db	601	ANAGCCCATTTGCNCTTTTTCNMANAGCCAAAGTTAAACCTGAANCTMTGCAATCTGG	660
OY	848	CTGCAGCTCTGACA	861
Db	661	GTGNACTTTTACCA	674

RESULT	12
LOCUS	AXI93464 518 bp DNA linear PAT 15-AUG-2001
DEFINITION	Sequence 1031 from Patent WO0149716.
ACCESSION	AXI93464
VERSION	AXI93464..1 GI:15211415
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 518)
AUTHORS	Xu,J., Lodes,M.J., Sectist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y. Compounds for immunotherapy and diagnosis of colon cancer and methods for their use Patent: WO 0149716-A 1031 12-JUL-2001;
JOURNAL	CORIXA CORPORATION (US)
FEATURES	location/Qualifiers source 1..518 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	123 a 136 c 149 g 109 t 1 others
ORIGIN	
Query Match	43.4%; Score 515; DB 6; Length 518;
Best Local Similarity	99.4%; Pred. NO. 6.1e-112;
Matches 515; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
OY	200 CCTGGGTGTTGAGGAATGGGCCCATTCACCAGGATCCTGGAGCGTGGTTGCAACATCC 259
DB	1 CCTGGGTGTTGAGGAATGGGCCCATTCACCAGGATCCTGGAGCGTGGTTGCAACATCC 60
OY	260 TCATCCCTTGTTAAACGCGATCCGATATGTGCAGAGTCTCAAGGAAATTGCATCAACG 319
DB	61 TCATCCCTTGTTAAACGCGATCCGATATGTGCAGAGTCTCAAGGAAATTGCATCAACG 120
OY	320 TGCCTGAGCAGTGGCGCTGTGACTCTTCGACAATGTAACTCTGCAAAATCGATGAGTCCCTT 379
DB	121 TGCCTGAGCAGTGGCGCTGTGACTCTTCGACAATGTAACTCTGCAAAATCGATGAGTCCCTT 180
OY	380 ACCTGGCATTATGACCTTTACAAAGGCAAGCTAGCGTGTGGAGAACCTTAGATATGCCG 439
DB	181 ACCTGGCATTATGACCTTTACAAAGGCAAGCTAGCGTGTGGAGAACCTTAGATATGCCG 240
OY	440 TCACCCAGGTAGCTCAAAACCAACATGAGATCAGAGCTGGCAAACCTCTTNNGACAAGA 499
DB	241 TCACCCAGGTAGCTCAAAACCAACATGAGATCAGAGCTGGCAAACCTCTCTTGACACAAG 300
OY	500 TCTTCGGGAAAGGAGATCCCTGAATGCCAGCATTTGTGATCCCATCAACCAAGTGTG 559

[illegible]

RESULT	13
AX339946	
LOCUS	AX339946
DEFINITION	Sequence 193 from Patent WO0196388.
ACCESSION	AX339946
VERSION	AX339946.1 GI:18135927
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (sites)
AUTHORS	Jiang,Y., Harlocker,S.D. and Secrist,H.
TITLE	Compositions and methods for the diagnosis of colon cancer
JOURNAL	Patent: WO 0196388-A 193 20-DEC-2001;
	CORIXA CORPORATION (US)
FEATURES	Location/Qualifiers
source	1..457 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	122 a 123 c 137 g 73 t 2 others
ORIGIN	

Query Match	38.3%	Score 455.2	DB 6	Length 457
Best Local Similarity	99.6%	Pred. No. 1e-97		
Matches 455	Conservative	1	Mismatches 1	Indels 0
				Gaps 0
QY	633 GCAGATGACAGGTGAGAGCGAGCGGCGGAACGGGGCCACAGTGTATAGAGTCTGAGGGGAC	692	1 GCAGATGACAGGTGAGAGCGAGCGGCGGAACGGGGCCACAGTGTATAGAGTCTGAGGGGAC	60
QY	693 CCGAGAGTCGGCCATCATGTGGCAGAGAGGAAGAACAGGCCACAGATCTGGGCTCCGA	752	693 CCGAGAGTCGGCCATCATGTGGCAGAGAGGAAGAACAGGCCACAGATCTGGGCTCCGA	752
Db	61 CCGAGAGTCGGCCATCATGTGGCAGAGAGGAAGAACAGGCCACAGATCTGGGCTCCGA	120	61 CCGAGAGTCGGCCATCATGTGGCAGAGAGGAAGAACAGGCCACAGATCTGGGCTCCGA	120
QY	753 AGCAGAAAAGGCTGAACACATTAATACAGCAGCAGAGAGGCGCAGTCTGGCGAA	812	753 AGCAGAAAAGGCTGAACACATTAATACAGCAGCAGAGAGGCGCAGTCTGGCGAA	180
Db	121 AGCAGAAAAGGCTGAACACATTAATACAGCAGCAGAGAGGCGCAGTCTGGCGAA	180	121 AGCAGAAAAGGCTGAACACATTAATACAGCAGCAGAGAGGCGCAGTCTGGCGAA	180
QY	813 GGGCAAGGCTAAAGCTGAGACTATTCTGAATCTGGCTGAGCTCTGACACAACATTAATGG	872	813 GGGCAAGGCTAAAGCTGAGACTATTCTGAATCTGGCTGAGCTCTGACACAACATTAATGG	872
Db	181 GGGCAAGGCTAAAGCTGAGACTATTCTGAATCTGGCTGAGCTCTGACACAACATTAATGG	240	181 GGGCAAGGCTAAAGCTGAGACTATTCTGAATCTGGCTGAGCTCTGACACAACATTAATGG	240
QY	873 AGATGAGACGCTTCACTGACAGTGGCGGAGAGATGTCACAGGGGTTCTCCAAACTGGC	932	873 AGATGAGACGCTTCACTGACAGTGGCGGAGAGATGTCACAGGGGTTCTCCAAACTGGC	932
Db	241 AGATGAGACGCTTCACTGACAGTGGCGGAGAGATGTCACAGGGGTTCTCCAAACTGGC	300	241 AGATGAGACGCTTCACTGACAGTGGCGGAGAGATGTCACAGGGGTTCTCCAAACTGGC	300
QY	933 CAAGGACTCCAAACACATCTCTACTGCGCTCCAAACCTGGCGATGTACACAGCANTGGTGGC	992	933 CAAGGACTCCAAACACATCTCTACTGCGCTCCAAACCTGGCGATGTACACAGCANTGGTGGC	992
Db	301 CAAGGACTCCAAACACATCTCTACTGCGCTCCAAACCTGGCGATGTACACAGCANTGGTGGC	360	301 CAAGGACTCCAAACACATCTCTACTGCGCTCCAAACCTGGCGATGTACACAGCANTGGTGGC	360
QY	993 TCAGGCCATGGGTGTATATGAGACCTCTCACCAAAGCCCATGTCACAGGAGACTCCAGACTC	1052	993 TCAGGCCATGGGTGTATATGAGACCTCTCACCAAAGCCCATGTCACAGGAGACTCCAGACTC	1052
Db	361 TCAGGCCATGGGTGTATATGAGACCTCTCACCAAAGCCCATGTCACAGGAGACTCCAGACTC	420	361 TCAGGCCATGGGTGTATATGAGACCTCTCACCAAAGCCCATGTCACAGGAGACTCCAGACTC	420

XX Goli SK, Hillman JL;
 XX MPI: 1998-347418/30.
 DR P-PSDB: AAM57232.
 XX
 PT DNA encoding human Integral membrane protein - useful for producing
 PT recombinant protein, for treatment of anaemia and cancer
 XX
 PS Claim 3: Column 37-40; 33pp; English.
 XX
 CC The present sequence encodes human Integral membrane protein (IMP).
 CC IMP may be administered to a subject to treat disorders associated
 CC with abnormal ion transport or membrane conductance as well as a
 CC variety of tumours, e.g. haemolytic anaemias and prostate, breast and
 CC pancreatic tumours. A vector capable of expressing IMP, or a fragment
 CC or a derivative thereof, may also be administered to a subject to treat
 CC the haemolytic anaemias and prostate, breast and pancreatic tumours.
 XX
 SQ Sequence 1188 BP: 268 A; 316 C; 362 G; 237 T; 5 other:

Query Match 99.6%; Score 1183; DB 19; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTGTGGAGNACGCGCTCGCTGCTGTTGTTCCGAGAGTCGCTCGCGGTGG 60
 Db 1 ggccttcggagcnaccgcctccgcctcgtcgttcgttcgaggtcgcgcgcgtcgtg 60
 QY 61 GAATGCTGGCG 120
 Db 61 gaaatgctggcg 120
 QY 121 GGCCTTGTGGCG 180
 Db 121 ggccttcggcg 180
 QY 181 TTCGTGCGCGAGAGAGCGCTGGGTGGTGGAGGAAATGGCGCATTCACCGATC 240
 Db 181 ttcgtgcgcgagagagcgctgggtgggtggaggaatggcgcattcacggatcctg 240
 QY 241 GAGCCTGGTTGAAACATTCCTATCCCTGTGTAGACGGATCCGATATGTGACAG 300
 Db 241 gagcctgggttgaacatctcctatccctgtgttagacggatccgatatgtgacag 300
 QY 301 AAGCAATTTGATCAACAGTGTGAGAGTGGGTGATCTCGAATGCAATGTAAC 360
 Db 301 aagcaatttgatcaacagtgtgagagtgggtgattctcgaaatgcaatgttaac 360
 QY 361 CAATGATGAGTCTTTTACCTCGCATGACCCCTTACAGGCAAGCTACGCTG 420
 Db 361 caatgatgagtccttttacctcgtcgtacatggaaccttaaggcaagctacgctg 420
 QY 421 GAGAGCCCTGAGTATGCGCTACCCAGTACTCAACAAACATGAGTACAGACTG 480
 Db 421 gagagccctgagtatgctgctacccagtactcaacaacatgagatcagagctg 480
 QY 481 AAATCTCTNTGACAAAGTCTTCCGGAAGGAGTCCCTGAATGCGAGATTGTG 540
 Db 481 aaatctctntgacaaagtcttccggaaaggagttccctgaatgcgagattgtg 540
 QY 541 GCCATCAACAGTGTGACTGCTGGGGTATCCGCTGCCCTTATGATCAAGAT 600
 Db 541 gccatcaacagtgtgactgctggggatccgctgcccttattgatcaagat 600
 QY 601 ATCATGATGCCACCGCGGTGAAGAGTCTATGAGATGCAAGTGGAGGAGAGCG 660
 Db 601 atcatgatgccaccggtgaagagtctatgagatgcaagtggagagagcgagcg 660
 QY 661 AAGGGGCAACAGTTTATGAGTGTGAGGAGACCGAGAGTGGCGCATATGTG 720
 Db 661 aaggggcaacagtttatgagtgtgagggagacccgagagagtgccaatgtgca 720

QY 721 GGGAAGAAACAGCCCAAGATCTGGCCTCCGAACGAGAAAAGCTGAACATTAATCAG 780
 Db 721 gggaagaagaaacagcccaagatctggcctccgaacgagaaagagctgacagataatcag 780
 QY 781 GCACGAGAGAGGCGCAGTTCGATTCGCGCAAGGCAAGGCTAAAGCTGAAGCTATT 840
 Db 781 gcacgagagagggcgagtcgagttctgcgcgaagggcaaggctaaagctgaagctattcga 840
 QY 841 ATCTGCTGAGCTGTGACATTAATGAGATGCGAGCTGCTGACTGTGGCC 900
 Db 841 atctgctgagctgtgacattaatgagatgcgagctgctgactgtggcc 900
 QY 901 GAGCAGTATGTCACGCGCTTCTCCAACTGGCCAGAGACTCAACACTATCTACTGCC 960
 Db 901 gagcagtatgtcacgcgcttctccaaactggccaaagactcaaacactactactgcc 960
 QY 961 TCACACCTGGCGCATGTCACAGCATGTGCTCAGCGCATGGGTGTTATGAGCCCTC 1020
 Db 961 tcacacctggcgcatgtcacagcatgtgctcagcgcatgggtgttatgagccctc 1020
 QY 1021 ACCAAGCGCCGAGTGCAGGAGCTCCAGCTCACTCTCAGTGGAGAGAGATGTC 1080
 Db 1021 accaagcccgagtgccagagctccagctcactctcagtgaggagagagatgctc 1080
 QY 1081 CAGGCTACAGATGCAAGTNTTATGAGAACTGATCGATGATCAAGATGATGAGGC 1140
 Db 1081 caggctacagatgcaagntnttatgagaaactgattcgatgacagatgagtgagtc 1140
 QY 1141 TGGGCTTNGCCAGGAGTCTGCGGAGCAAGCAAGATTTTCTGAT 1188
 Db 1141 tgggcttngccaggagtctcgaggagagagatlttccctgatt 1188

RESULT 2
 AAX04402
 ID AAX04402 standard; DNA: 1322 BP.
 AC AAX04402;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Human secreted protein gene 35 clone HTX521.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 XX
 PN W03856804-A1.
 PD 17-DEC-1998.
 XX
 PF 11-JUN-1998; 98MO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.

OY 960 CTCACACCTGGGATGTCACAGATGTTGGTGGCCAGGCGCATATATGGAGCCCT 1019
|||||
DB 434 CTCACACCTGGGATGTCACAGATGTTGGTGGCCAGGCGCATATATGGAGCCCT 375
|||||
OY 1020 CACCAAGGCCAGTGGCAGGAGCTCCAGACTCTCCAGTGGGAGCAGAGATGT 1079
|||||
DB 374 CACCAAGGCCAGTGGCAGGAGCTCCAGACTCTCCAGTGGGAGCAGAGATGT 315
|||||
OY 1080 CCAGGCTACGATGCAAGTATTGATGAGAACTTGATCGATCAAGATGATTAGTGAG 1139
|||||
DB 314 CCAGGCTACGATGCAAGTATTGATGAGAACTTGATCGATCAAGATGATTAGTGAG 255
|||||
OY 1140 CTGGGCTTGGCCAGGAGCTGGGGACAAAGAGCAGATTTTCTGATT 1188
|||||
DB 254 CTGGGCTTGGCCAGGAGCTGGGGACAAAGAGCAGATTTTCTGATT 206
|||||
RESULT 4
AAK52683/c
ID AAK52683 standard; cDNA: 1398 BP.
XX AAK52683;
AC
XX
XX
DT 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 2212.
DE
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
XX
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX
XX nervous system disorder; arthritis; inflammation; ss.
OS
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX
XX 20-JUN-2000; 2000US-0598075.
XX
XX 19-JUL-2000; 2000US-0620325.
XX
XX 01-SEP-2000; 2000US-0654936.
XX
XX 15-SEP-2000; 2000US-0663561.
XX
XX 20-OCT-2000; 2000US-0693325.
XX
XX 30-NOV-2000; 2000US-0728422.
XX
XX
XX (HYSE-) HYSEQ INC.
XX
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX
XX WPI: 2001-476283/51.
XX
XX P-PSDB: AAM79550.
XX
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX
XX useful in diagnosis and gene therapy -
XX
XX
XX Claim 1: Page 4572: 6221pp; English.
XX
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX
XX cytokine, cell proliferation or cell differentiation or which may induce
XX
XX production of other cytokines in other cell populations. The
XX
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX
XX activity, tissue growth factor activity, immunomodulatory activity and
XX
XX activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC NOTE: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX

SQ Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;

Query Match 97.6%; Score 1160; DB 22; Length 1398;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

OY 1 GGCCTTGGGAGGACACCGCTCCGCTGCTCTGTTGTTCCGGAGAGTGGCGGGCTGG 60
|||||
DB 1393 GGCCTTGGGAGGACACCGCTCCGCTGCTCTGTTGTTCCGGAGAGTGGCGGGCTGG 1334
|||||
OY 61 GAATGCTGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
|||||
DB 1333 GAATGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1275
|||||
OY 121 GGCCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
|||||
DB 1274 GGCCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1215
|||||
OY 180 GTTCGTCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 239
|||||
DB 1214 GTTCGTCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1155
|||||
OY 240 GAGAGCCGTTGTTGAAACATCCCTCATCCCTGTTAGACCGGATCCGATATGCGAGTCT 299
|||||
DB 1154 GAGAGCCGTTGTTGAAACATCCCTCATCCCTGTTAGACCGGATCCGATATGCGAGTCT 1095
|||||
OY 300 CAAGGAATTTGTCATCAGCTGCTGAGCAGTGGCTGTGACTGTGACAATGTAATCT 359
|||||
DB 1094 CAAGGAATTTGTCATCAGCTGCTGAGCAGTGGCTGTGACTGTGACAATGTAATCT 1035
|||||
OY 360 GCAATGATGAGAGTCCCTTACCTGCGCATCATGAGACCTTACAAAGCAAGCTACGCTGT 419
|||||
DB 1034 GCAATGATGAGAGTCCCTTACCTGCGCATCATGAGACCTTACAAAGCAAGCTACGCTGT 975
|||||
OY 420 GAGGAGCCCTGAGATATGCGCTCAACCCAGTACGTCAAACCAACATGAGATCAGAGCTCG 479
|||||
DB 974 GAGGAGCCCTGAGATATGCGCTCAACCCAGTACGTCAAACCAACATGAGATCAGAGCTCG 915
|||||
OY 480 CAAACTCTTNTGACAAAGCTTTCGCGGGAACGAGAGTCCCTGATATCCAGATTTGGA 539
|||||
DB 914 CAAACTCTTNTGACAAAGCTTTCGCGGGAACGAGAGTCCCTGATATCCAGATTTGGA 855
|||||
OY 540 TGGCATCAACCAAGCTGCTGAGTGGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
|||||
DB 854 TGGCATCAACCAAGCTGCTGAGTGGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
|||||
OY 600 TATTCATGTCACACCGCGGCTGAAGAGTATGCAATGATGAGTGGAGCAGAGCGCG 659
|||||
DB 794 TATTCATGTCACACCGCGGCTGAAGAGTATGCAATGATGAGTGGAGCAGAGCGCGCG 735
|||||
OY 660 GAAACGGGCCACAGTTCTAGAGTCTGAGGGACCCGAGAGTGGGCCATCAATGTGCGAGA 719
|||||
DB 734 GAAACGGGCCACAGTTCTAGAGTCTGAGGGACCCGAGAGTGGGCCATCAATGTGCGAGA 675
|||||
OY 720 AGGGAAGAAACAGGCGCCAGATTCCTGGGCTCGAAGCAAGAAAGGCTAACAAGATAATCA 779
|||||
DB 674 AGGGAAGAAACAGGCGCCAGATTCCTGGGCTCGAAGCAAGAAAGGCTAACAAGATAATCA 615
|||||
OY 780 GGCAGCAGAGAGGCGCAGTCAAGTTCGCGGAAGGCCAAGGCTAAAGCTAAGCTATTTCG 839
|||||
DB 614 GGCAGCAGAGAGGCGCAGTCAAGTTCGCGGAAGGCCAAGGCTAAAGCTAAGCTATTTCG 555
|||||
OY 840 AATCTGGCTGACGCTTGACACAAATATGAGATGACAGAGCTTCACTGACTGGCG 899
|||||
DB 554 AATCTGGCTGACGCTTGACACAAATATGAGATGACAGAGCTTCACTGACTGGCG 495
|||||

OY	495	CAAGTCTTCCGGGACAGGGAATGCTTGAATGCCAGCATTTTGATGCCATTCACCAACG	554
Db	507	caaatcttcctcgggaaacggyagtccttgaatgcacgacatlttgagatgcatcaacaaacg	566
OY	555	TGCTGACGCGGGGGTTCGCGTCGCGMNGGTTATGAGATCAAGGAAATTCATGTTCCACC	614
Db	567	ttcgcacatgcgcggggatctaccgcctgcgccttctatgagatacaagatatccatgtgcacc	626
OY	615	CCGGGTCAAGAGACTATATGCAGATGCAGGTGGAGGAGCAGCAGCGCGGAAACGGGCACAGT	674
Db	627	ccgggtgaaagagtctctatgagatgcaagtgtgagggcgagggcggygaaacgycacacgt	686
OY	675	TCTAGAGCTTGAGGGGACCCGAGAGTGGCCATCAATGTGGCAGAGGAGGAAAGAAACAGC	734
Db	687	tctaagatcttgaaggagcccgagagtcggccatcaatgtgagagaaggaagaacacagc	746
OY	735	CCAATCTCTGGCTCCGAAACAGAAAAGCTGTAACAGATAAATCAGGCAGCAGAGAGGC	794
Db	747	ccaaatctctgcccctccgaagcagaanaagcctgaacagataaaatcagcagcagagagagc	806
OY	795	CAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAGCTGAAGTATTTCAGATTCCTGGCTGCAGC	854
Db	807	cagtgccagttctcggcgyaagcccaagcctlaaagctgaagctatctgaaatccctgycgcagc	866
OY	855	TCTGACACAACATATATGAGATGCACACAGCTTCACCTGATGTGGCGGAGCAGTATGTAG	914
Db	867	ctctgacacaacataatgagagatgcagcagcttcacgcactgtgycgcgagcagatgatctag	926
OY	915	CGCGTTCTCCAAACTGGCCCAAGGACTTCACACACTATCTACTGCGCTTCACACCTGGCGA	974
Db	927	cgcgcttccaaactgyccaagactccaacatactcactacgtccctccaaacctgcyga	986
OY	975	TGTACACAGATGGTGGCTATGAGGCCATGGGTGATATATGAGGCCCTCACCAAGCCCCAGT	1034
Db	987	tgtaaccagaatggtgtgtctcaaggccatggtgtgtatatgtgagccctccacaagaagcccgct	1046
OY	1035	GCCAGGACATCCAGACACTCTCCAGTGGAGACAGACAGATGTTCCAGGGTACAGATGC	1094
Db	1047	gccagggactcccaactcactctcccaagtggagcagcagaagatgtgccaggttacaagatgc	1106
OY	1095	AAGTTTGTATGAGAACTTGATCGAGTCAAGATGAGTTAGTGGAGCTGGGCTTNGCCAGG	1154
Db	1107	aagctctgatgagaaactgtatcgcagtcgaatgagttagtgagagcttgyccttgccag	1166
OY	1155	GAGTCTGGGACAAAGAGACAGATTTTCTGATT	1188
Db	1167	gagctctgggacaagaagacagatlttctccgatt	1200

RESULT	6	
AAK51699		
ID	AAK51699	standard; cDNA: 1416 BP.
XX		
XX		
AC	AAK51699;	
XX		
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human polynucleotide SEQ	ID NO 244.
XX		
KW	Human: cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation; ss.	
XX		
OS	Homo sapiens.	
XX		
XX		
PN	WO200157190-A2.	
XX		
XX		
PD	09-AUG-2001.	
XX		
XX		
PF	05-FEB-2001; 2001WO-US04098.	
XX		

PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

DR WPI; 2001-476283/51.
DR P-PSDB; AAM78566.

PT Nucleic acids encoding polypeptides with cytokine-like activities
PT useful in diagnosis and gene therapy -

PS Claim 1; Page 1135-1136; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK5343) and the encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activating/inhibiting activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

SQ Sequence 1416 BP; 335 A; 364 C; 410 G; 307 T; 0 other;

Query Match 96.38; Score 1144; DB 22; Length 1416;

Best Local Similarity 98.98; Pred. No. 8.1e-315;

Matches	1182;	Conservative	0;	Mismatches	5;	Indels	8;	Gaps	3;
---------	-------	--------------	----	------------	----	--------	----	------	----

QY 1 GGCTCTGGGAGCNAACGGCTCCGCTCGTCTGTTGGTTCGGAGGTCCCTGGCGCGGTG
 |||||
 Db 11 qgcctctggaagcaaccgcctccgctcgtctgcttgcgcgaagctcgcctgcgcgaagtgc

61 GAAATGCTGGCCGCCCGGCAGGCGTCCTCA

0.1 GAAAGTCTGGCGGCGCGCGCACATGGGCCCTTTTCCTGAAGGGCCTCCAC

27

Dh 71 aaaaataaacacacacacacact'aaacacactt'actaaacacactctctac

[illegible][illegible]

121 GGCTTCCTGGCCGGCTCCG-GCCGGCCCTCCCTCCTGGATTGCCCGAAACACCGTGGTAC

[illegible]

Db 130 ggctctcgcgcgcctccgcgcgcctcctggaattgccccgaaacacccgtgtac

[illegible]

180 GTTCGTGCCGACGAGGCCCTGGGTGGTGGAGCGAATGGGCCGATCCACCGGATCC

Db 190 gttcgtgccgcaagagcctgtgtgtgagcgaatgtggccgattccaacggatcc

.....

240 GGAGCCTGGTTTGACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCT

Db 250 ggagcctggttgacatcctcatccctgtgttagaccggtatccgatatgtgcagagtc

QY 300 CAAGGAATTGTCATCAACGTGCCCTGAGCAGTCCGCTGTGACTCTCGACAATGTA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 310 caaggaattgtcatcaacgtgctgagcagtcgctgtactctcgacaatgtaactct

.....

360 GCAATCGATGGAGTCCCTTACCTGCCCATCATGGACCCCTACAAGGCCAAGCTACGGTGT

100

Db 370 acaaatcgaatggaatcctttacctaagcatcatgaaccccttacaaggcaaaactacggtat

[illegible]

```

QY 420 GGAGACCCCTGAGTATGCTGCTACCCAGCTACGTCACCAACCATGATGACAGCTCGG 479
    |||
Db 430 ggaagaaacctgagtaatgacgctacccagctagctcaaaacacagatcagagctcg 489
QY 480 CAACCTCTCTTGTGCAAGAGTCTCCGGGACCGGAGTCCCGATGCGCATGTGTGGA 539
    |||
Db 490 caaacctctctcggcaaaagctctccggaaacggagctcccgaaatccagacatctgga 549
QY 540 TGCCATCAACCAACGCTGCTGATGCTGGGGTATCCGCTGCTTCGCTTATGAGATCAAGA 599
    |||
Db 550 tgcacatacaacagctgctgactgctggtgatactcgcctccgtctatagataaagga 609
QY 600 TATCCATGTCACCCCGGGGTAAAGACTCTATGACAGATGCAAGTGAGGCGAGCGCGC 659
    |||
Db 610 tatccatgtgcacaccccggtgaaagatctatacgagatgcagtggaagcgagcgcg 669
QY 660 GAAACGGGCGACAGTTTGAAGTCTGAGGGGACCGGAGAGTGGGCGCATGAATGTGACAGA 719
    |||
Db 670 gaaacgggccaagatctcagagatctggaagggaacccgagagctcgacataatgtggcaga 729
QY 720 AGGGAAGAAACAGGCCCAAGATCTGCTGCTCCGAGACAGAAAAGGCTGAACAGATTAATCA 779
    |||
Db 730 agggagaagaaacagcgccagatctcgtgctccgaaagcagaagagctgaaacagataatca 789
QY 780 GGCACAGAGACAGCGCAGTGTCTGGCCGAAGCCCAAGGCTAAAGCTGAAGCTATTGG 839
    |||
Db 790 ggcacagagagagcgacgtgctgtggaaggcgaagctgaaagctgaaagctatctg 849
QY 840 ATCTGCTGCTGAGCTCTATACACATATATGAGATGAGAGAGTTCATGACTGTGCG 899
    |||
Db 850 aatctgtgctgacgtctgacacacataatgagagatgcagagctctcgtgactggtgc 909
QY 900 CGAGAGATATGTACAGCGCTTCTCCAAACTGGCCCAAGAGACTCCAACTATCTACTGCG 959
    |||
Db 910 cgagagatgtgcagcgctctctcaaaactgccaagagctccaacacatctactgac 969
QY 960 CTTCACACCTGGCGATGTCACAGATGCTGGCTCAGGCGCATGGGGTATATGAGACCTT 1019
    |||
Db 970 ctccaacccctggagatgtcacacagatgtgctcagcgccatggtgatalatggaagccct 1029
QY 1020 CACCAAGC-----CCAGTGCACAGGAGCTCCAGACTCTCTCCAGTGGGAGCAGCAG 1073
    |||
Db 1030 caccaaagccccaagctccagtcagtcagagagctccagctcactctccagtcggaagcgag 1089
QY 1074 AGATGTCAGGCTACAGATGCAAGTATTGATGAGAACTTGTGATCAAGTGAAGTGA 1133
    |||
Db 1090 agatgtccaggggtacagatgcagatctgtatgaggaactgtacgagtcacagatgagtt 1149
QY 1134 GTGGAGCTGGGCTTNGCCAGGAGTCTGGGAGCAAGAGCAGATTTCCTGATT 1188
    |||
Db 1150 gtggagctgtggctgtgcccagggagctctggtggaagaaagcagatcttccctgatt 1204

```

RESULT 7
AAH13961
ID AAH13961 standard; cDNA: 1244 BP.

AAH13961:

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:11009.

Human: primer: detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

```

PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI: 2001-318749/34.
XX
XX Claim 8: SEQ ID 11009; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC of the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH0246 to
CC AAH05893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1244 BP; 284 A; 327 C; 368 G; 265 T; 0 other;

```

Query Match 95.2%; Score 1130.4; DB 22; Length 1244;
Best Local Similarity 99.4%; Pred. No. 66-311;
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```

QY 30 TCGTTGGTTCGGAGAGCTCGCTGCGGCGGTGGGAAATGCTGCGCGCGCGGCGGGGCGCA 89
    |||
Db 1 tcgtgtgttcggaagtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 59
QY 90 CTGGGCGCTTTTGTGAGAGGCGCTCTCTACTGCTTGTGCGCGCTCGG-CGCGGCGCT 148
    |||
Db 60 ctggggcccttctgctggaagggctctctactgctctgctgctgctgctgctgctgctgct 119
QY 149 CCTCTGATTTGCCCGCAAAACACCGTGTACTGTTGCTGCCGACGACGAGAGCGCTGGGTG 208
    |||
Db 120 cctctgattgcccgcaaacacgctgtgtactgtctgtgctgctgctgctgctgctgctgct 179
QY 209 TGGAGCGAATGGGCGCATTCACCGGATCTGAGAGCTGGTTTGAACATTCCTATCCCTG 268
    |||
Db 180 tggagcgaatgggcgatctcacccggaatctcgtgagctgtgttgaacatctcatccctcgt 239
QY 269 TGTTAGCCGAGATCCGATATGTGACAGTCTCAAGGAAATGTGCATCAACGTCCTGAGC 328
    |||
Db 240 tgttagaccggaatccgatatgtgcagagctcctaagaaatgtgcacaaagctgctcgagc 299
QY 329 AGTCGGCTGTGACTCTGACAAATGTAAGTGAATGCAATGCAATGCAATGCAATGCAATGCA 388
    |||
Db 300 agtcgctgtgactctcgaaatgtaactctgcaaatcgatgagtgatccttactctgca 359

```

QY 389 TCATGACCCCTTACAAAGCAGCTACGCTGTGAGAGACCCCTGATGCTGACCCAGC 448
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 360 tcattgacccttaacaaggaaagctacggtgtgagagacccttgatgctgcgcacaccagc 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 449 TAGCTCAAAACATCATGATGATGAGCTGCGCAAACTCTCTNTGCAAAATCTTCCGGG 508
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 420 tagctcaaaacacacatagatcagagctcgcaaacctctctctggaacaaagctctccgg 479
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 509 AACGGGAGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 480 aacgggagtccttgatgacagcatgtgagatgacacacacagctgctgacgtctggg 539
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 559 GTATCCGCTGCTTCGCTTATGATATCAAGATATCATGCTGCTGCTGCTGCTGCTGCTG 628
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 540 gatacgcgcctcgcctcgtatgagatcaagaatacctatgctgcaaccccggtgaaagat 599
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 629 CTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 600 ctatgagatgagatgtgagagcagagcggtgaaacgggccaagctctcagagctcagag 659
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 689 GGACCCGAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 660 ggaacccgagatgctgacatcaatgtgagagaggaagaaacagccagctcctggtcct 719
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 749 CCGAAGCAGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 720 ccgaagcagaaagagctgacagataaatacaagcagcagagagagagagagagagagag 779
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 809 CGAAGCCCAAGCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 868
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 780 cgaagcagagagcttaagctgagagctgagatcctgagctgagctcgcagacacacata 839
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 869 ATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 840 atggagatgacagagctcctacgagctggtgagagagagagagagagagagagagagag 899
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 929 TGGCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 988
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 900 tggcgaagagctcgaacatct 959
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 989 TGGCTCAGCCATGGGTGTATGTGAGAGCCCTTCAACAAAGCCCAAGTGCACGACTCCAG 1048
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 960 tggctcagagcctatgtgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 1019
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1049 ACTGCACTCTCCAGTGGGAGCAGACAGATGATGATGATGATGATGATGATGATGATGATG 1108
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1020 actcactctccagtggtgagcagagagatgtccaggtgacagatgcaagctctgagag 1079
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1109 AACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1168
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1080 aacttgatcagagtgagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1139
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 1169 GGAAGCAGATTTTCTGATT 1188
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1140 ggaagcagatttctcgtalt 1159
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 8

AAS44953 standard: cDNA: 9098 BP.
 AAS44953:
 18-DEC-2001 (first entry)
 DE cDNA encoding novel human secretory protein, Seq ID No 34.
 XX Human: secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KM Ischemia-reperfusion injury; haematopoiesis: cancer; neuropathy;
 KM transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KM amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KM ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KM gut protection: lung; liver fibrosis; immune deficiency; infection;
 KM severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KM multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KM fertility; analgesic; pain; antigen; ss.
 OS Homo sapiens.
 XX WO200166689-A2.
 PN 13-SEP-2001.
 PD 05-MAR-2001; 2001WO-US04942.
 PF 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0618647.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX (HSE-) HSEQ INC.
 PA Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Dirmanc RT, Zhang J, Chen R, Xue AJ, Wang J;
 PI WPI: 2001-589934/66.
 DR P-PSDB: AAN28053.
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 PT XX
 PS Claim 1: SEQ ID No 34; 107pp; English.
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAS44920-AAS45295 represent novel human secreted protein
 CC coding sequences of the invention.
 CC XX
 SQ Sequence 9098 BP; 1918 A; 2538 C; 2590 G; 2052 T; 0 other:

Query Match 93.0%; Score 1104.8; DB 22; Length 9098;
 Best Local Similarity 99.3%; Pred. No. 2.8e-303;

Matches 1128: Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```

Oy 1 GGCCTTGTGGAGACNACCGCTCCGTCGTCTGTTGGTTCGGAGGCTCGCTGGCGGCTGG 60
    |||
Db 12 ggctctctggagcgagacgctccgctcgtctcgtctcgtctcgtcggagcgcgcg 71
Oy 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGACATGGGCGCTTTTGGCGGGGCTCTTACT 120
    |||
Db 72 gaatctgctggcgcgcgcgcg-99ggacatggggcccttcgcggaagggtctctact 130
Oy 121 GCGCTTCTGGCCCCCGCTCCG-GGCGGCGCTCCCTCTGTGATTGGCCCCGAACACCGTGTACT 179
    |||
Db 131 ggctctggcgcgcgctcgcgcgcgcctcctcctcgtatggccccgaacacccgtgtact 190
Oy 180 GTTCTGCGCCGACGACGAGAGGCTGGGTGGTGGAGCAATGAGCGCGCTTTCACCGGATCT 239
    |||
Db 191 gctgtgctgcgcgcgcggggtcgtgggtggagcgaaatgggcccgtctccacggatcct 250
Oy 240 GGAGCCTTGTTGAACATCTCATCTCTGTGTAGACCGGATCCGATATGTGACAGTCT 299
    |||
Db 251 ggagcctgtgtgaacatccatccatccctgtgttagaccggatccglatgtgcagagct 310
Oy 300 CAGGAATTTGTCATCAACAGTGGCTGAGAGTGGCTGTGACTCTGACAAATGTACTCT 359
    |||
Db 311 caaggaaatgtcatcaacacgtgcttgagcagtcggctgtgactctcgacaatgtactct 370
Oy 360 CCAATCGATGAGTCTTCTTACCTGCGATCATGACACCTTACAAAGCAAGCTACGGTGT 419
    |||
Db 371 gcaatctgattgagctccttaccctgacatcattgaaaccttcgaaggcaagctaaaggtc 430
Oy 420 GGAGGACCTTGATATGCCGTGACCCAGCTAGCTCAAAACACATGAGATGAGAGCTCG 479
    |||
Db 431 ggaggacctgtglatgctgctacacagctagctcaaaaccaaagatcagagctcgg 490
Oy 480 CAAACTCTCTNTGCAAAAGTCTTCCGGGAAGGAGTCCCTGATGCTCCAGATTTGTGA 539
    |||
Db 491 caaactctctcggaaagatctccggaaacggagatcccgaaaggcagatctgttga 550
Oy 540 TGCCTCAACCAAGCTGCTGATGCTGGGTATCCGCTCTCCCTTATGATGATCAAGA 599
    |||
Db 551 tgcctcaaccaaagctgtcgtactgtgtgggtatccgctcctcgctatgataaagaa 610
Oy 600 TATCCATGTCCACCCCGCGCTGAAGAAGTCTATGCAAGATGACGTTGAGGACGAGCGCG 659
    |||
Db 611 tatccatgtgccaccccggtgtaaagatctatgcaagatgcaggtggagcgagcg 670
Oy 660 GAAAGGGGCGACAGTTCTAAGATCTGAGGGGACCCGAGAGTGGCGATCAATGTGGCA 719
    |||
Db 671 gaagcgccacagatctctagatctgaggggagcccgaggtcggcatalatgtggcaga 730
Oy 720 AGGGAAGAAACAGGCGCCAGATCTGGCTCCGACAGCAGAAAGCTGAGCAGATAAATCA 779
    |||
Db 731 agggagaaacagcgccagatctcgtccctcgagagcagaagaaagctggaacagataatca 790
Oy 780 GGCACGACGAGAGGCGCAGTGTCTGGCGAAGGCCAAAGCTAAAGCTGAAGCTATTTC 839
    |||
Db 791 ggacagcagagagggcaggtgcagttctggcgaagccaaggtcaaggtcgaagctatcgc 850
Oy 840 AATCTGGCTGCAGCTCTGACACACATTAATGAGATGACAGAGCTTCACTACTCTGGC 899
    |||
Db 851 aatctgtgctgagatctcgaacacataatgagatgcgaagatctcctcgtactgtgc 910
Oy 900 CGAGCAGTATGACAGCGCTTCTCAAACTGGCAAGAGCTCAAGACTATCTCTACTGGC 959
    |||
Db 911 cgaagcagatgtcagcggtctctccaaactggcgaagatctcaacgctcatctctatcgc 970
Oy 960 CTGCAACCTTGCGATGTCTACAGCATGGTGGCTCAGGCCATGGCTGATATGAGCCCT 1019
    |||
Db 971 ctccaacctgtgagatgtcaccagcatgtggctcagggcatalgtatataatggccct 1030
Oy 1020 CACCAAAACCCCAAGTGCAGGAGACTTCAGACTCTTCTCAATGGGAGCAGCAGAGATGT 1079
    |||
Db 1031 caccaaaagcccaagtgccagaggtcccgatctcactctccagtggtggcagcagagatgt 1090

```

```

Oy 1080 CCAGGTACAGATGCAGATNTGATGAGAACTTGATCGAGTGAATGATTACT 1135
    |||
Db 1091 ccagggtacagatgcagatctctgtatgtaggaactgtatcgtacaaatgacttgat 1146

```

RESULT 9

AAS58752
ID AAS58752 standard; cDNA; 691 BP.

AC AAS58752;

DT 13-FEB-2002 (first entry)

DE cDNA #1428 encoding portion of a human colon tumour protein.

KM Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.

XX Homo sapiens.

PN W0200173027-A2.

PD 04-OCT-2001.

PF 22-MAR-2001; 2001WO-US09246.

PR 24-MAR-2000; 2000US-191597P.

PR 04-MAY-2000; 2000US-202024P.

PR 05-MAY-2000; 2000US-202189P.

PA (CORI-) CORIXA CORP.

PI Meagher MJ, Xu J, King GE;

DR WPI: 2001-611627/70.

PT New colon tumour proteins and related nucleic acid, useful for

PS treatment, prevention, diagnosis and monitoring of cancer.

XX Claim 4; Page 281; 299pp; English.

CC Th present invention relates to the isolation of novel cDNA sequences
 CC encoding for at least an immunogenic portion of human colon tumour
 CC proteins. The sequences of the invention are useful in pharmaceutical
 CC compositions and vaccines for the prevention and treatment of cancers
 CC such as colon cancer. They are also useful for the diagnosis and
 CC monitoring of such cancers. Antibodies to the colon tumour proteins
 CC and antigen presenting cells that express polynucleotides encoding
 CC colon tumour proteins can be used to inhibit the development of
 CC cancers. T-cells that react specifically with colon tumour proteins
 CC are useful for removing tumour cells from samples (e.g. blood) and
 CC for cancer treatment. The polynucleotide sequences are also useful in
 CC gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
 CC invention that encode for portions of human colon tumour proteins.

XX Sequence 691 BP; 183 A; 168 C; 184 G; 140 T; 16 other;

Query Match 43.5%; Score 516.6; DB 23; Length 691;
 Best Local Similarity 91.4%; Pred. No. 1.2e-136;
 Matches 616; Conservative 0; Mismatches 46; Indels 12; Gaps 7;

```

Oy 200 CCGTGGTGTGGAGGAATGGGCGGATTCACCGGATCCCTGAGCGCTGTTGAACATCC 259
    |||
Db 1 cctgggtgtgtgaggaatgggcgcatccacggatccctgagcctgtgtgaaacatcc 60
Oy 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGAAATTTGTCATCAAG 319
    |||
Db 61 tcatccctgtgttagaccggatccgatattgtgcagagctccaaggaattgtcatcaacg 120
Oy 320 TCCCTGAGCAGTGGCGCTGTGACTCTCGACATGTAACTCTGCAATTCGATGGAGCTTTT 379
    |||
Db 121 tgcctgagcagtcgctgtgactctcgacaatgtactctgcaaatctgtatgtggatccttt 180

```


• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:26:15 ; Search time 50.04 Seconds
(without alignments)
5831.587 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GCGTCTGCGAGCACCCTGCT.....GGAAGCAGATTTCGATT 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	99.6	1188	1	US-08-781-562-2
2	48.8	4.1	7218	1	US-08-232-463-14
3	39.6	3.3	2488	1	US-08-279-270A-2
4	39.2	3.3	2557	4	US-08-464-954A-1
5	39	3.3	2204	1	US-08-221-817-12
6	39	3.3	2204	1	US-08-454-439-12
7	39	3.3	2204	5	PCT-US94-10487-12
8	39	3.3	2206	1	US-08-221-817-10
9	39	3.3	2206	5	US-08-454-439-10
10	39	3.3	2206	5	PCT-US94-10487-10
11	39	3.3	7791	2	US-08-149-097D-23
12	39	3.3	7791	3	US-08-949-386-23
13	39	3.3	7791	3	US-08-450-562-23
14	39	3.3	7791	3	US-08-984-709A-23
15	39	3.3	7808	4	US-08-149-097D-22
16	39	3.3	7808	3	US-08-949-386-22
17	39	3.3	7808	3	US-08-450-562-22
18	39	3.3	7808	4	US-08-984-709A-22
19	37.6	3.2	289	4	US-09-007-005-17
20	37.6	3.2	289	4	US-09-244-796-17
21	37.4	3.1	2848	4	US-08-464-954A-2
22	37.2	3.1	977	6	5215895-2
23	36.2	3.0	1983	1	US-08-221-817-21
24	36.2	3.0	1983	1	US-08-454-439-21
25	36.2	3.0	1983	5	PCT-US94-10487-21
26	36.2	3.0	3224	5	US-09-079-415-3
27	35.6	3.0	977	1	US-08-017-522A-1

28	35.6	3.0	1100	1	US-07-949-516A-1	Sequence 1, Appli
29	35.6	3.0	1100	2	US-08-814-459-1	Sequence 1, Appli
30	35.6	3.0	1100	3	US-09-122-525-1	Sequence 1, Appli
31	35.2	3.0	13842	4	US-09-105-537-30	Sequence 30, Appli
32	35.2	3.0	36778	4	US-09-105-537-5	Sequence 5, Appli
33	35.2	3.0	38506	3	US-09-320-878-19	Sequence 19, Appli
34	35	2.9	500	3	US-09-141-000-2	Sequence 2, Appli
35	34.8	2.9	2018	2	US-08-557-973-1	Sequence 1, Appli
36	33.8	2.8	2277	1	US-08-676-967-2	Sequence 2, Appli
37	33.8	2.8	2277	1	US-08-676-974-2	Sequence 2, Appli
38	33.8	2.8	2277	2	US-09-098-487-2	Sequence 2, Appli
39	33.4	2.8	8252	1	US-08-046-585-15	Sequence 15, Appli
40	33.4	2.8	8252	1	US-08-393-703-15	Sequence 2, Appli
41	33.4	2.8	8252	5	PCT-US93-11721-15	Sequence 15, Appli
42	33.2	2.8	2307	3	US-08-942-008-1	Sequence 1, Appli
43	32.8	2.8	1812	2	US-08-735-041A-3	Sequence 3, Appli
44	32.8	2.8	1812	3	US-09-190-476B-3	Sequence 3, Appli
45	32.8	2.8	1812	3	US-09-190-889A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-781-562-2
; Sequence 2, Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,562
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PR-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1188 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-781-562-2
Query Match 99.6%; Score 1183; DB 1; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	GGATTGCGGAGGACNACGGCTCCGGCTCTGTCGTTGGTTCCGGAAGGTCCTGGCGGGTGG	60
Db	1	GGCTTCTGGGAGGACNACGGCTCCGGCTGTCCTGTGTTCCGGAGGTCCTGGCGGGTGG	60
QY	61	GAATTCGTGGCGGCGGCGGCGGAGGAGCTGGGGCCCTTTTGTCTGAGGGGCTCTACT	120
Db	61	GAATTCGTGGCGGCGGCGGCGGAGGAGCTGGGGCCCTTTTGTGAGGGGCTCTACT	120
QY	121	GGCTTCTGGCGGCGGCTCCGGGCGGCTCTCTGATTTGCCCGAACAAGCTGGTACTG	180
Db	121	GGCTTCTGGCGGCGGCTCCGGGCGGCTCTCTGATTTGCCCGAACAACGTTGGTACTG	180
QY	181	TTTCTCCCGCAGAGAGAGGCGCTGGGTGTGTGAGACGAATGGGCGGATTCACCGGATCTG	240
Db	181	TTTCTCCCGCAGAGAGAGGCGCTGGGTGTGTGAGACGAATGGGCGGATTCACCGGATCTG	240
QY	241	GAGCTGTGTTGAACATCTCTATCCCTGTGTTAAGACCGGATCCGATATGTGACAGTCTC	300
Db	241	GAGCTGTGTTGAACATCTCTATCCCTGTGTTAAGACCGGATCCGATATGTGACAGTCTC	300
QY	301	AAGGAATTTGTCAATCAACGTCGCTGTGAGACGATGGGGGTGATCTGTGACAATGTAACTGTG	360
Db	301	AAGGAATTTGTCAATCAACGTCGCTGTGAGACGATGGGGGTGATCTGTGACAATGTAACTGTG	360
QY	361	CAATTCGATGGAGTCCTTTACCTGCGCCATCATGAGACCTTTACAAAGGCAAGCTACGGTGTG	420
Db	361	CAATTCGATGGAGTCCTTTACCTGCGCCATCATGAGACCTTTACAAAGGCAAGCTACGGTGTG	420
QY	421	GAGGACCCCTGAGTATGCGCGTCACCCAGAGCTATGCTCAAAACACCATAGATCAGAGCTGGC	480
Db	421	GAGGACCCCTGAGTATGCGCGTCACCCAGAGCTATGCTCAAAACACCATAGATCAGAGCTGGC	480
QY	481	AAATCTCTTMTGACAAAGTCTTCCGGAACGGGAGTCCCTGAATGCCAGATTTGTGAT	540
Db	481	AAATCTCTTMTGACAAAGTCTTCCGGAACGGGAGTCCCTGAATGCCAGATTTGTGAT	540
QY	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCCTNCGTATAGATCAAGAT	600
Db	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCCTNCGTATAGATCAAGAT	600
QY	601	ATTCATTTGCGCACCCCGGGGGAANAAGTCTATGACGATGACAGTGGAGGACGAGCGGG	660
Db	601	ATTCATTTGCGCACCCCGGGGGAANAAGTCTATGACGATGACAGTGGAGGACGAGCGGG	660
QY	661	AAACGGGCGCACAGTTCTTAGAGTCTGAGGGGACCCGAGATCGGCCATCAATGTGGCAGAA	720
Db	661	AAACGGGCGCACAGTTCTTAGAGTCTGAGGGGACCCGAGATCGGCCATCAATGTGGCAGAA	720
QY	721	GGGGAAGAAACAGGCCCAAGTCCCTGGGCTCCGAAGCAGAAAGCGCTGAACAGTAAATAG	780
Db	721	GGGGAAGAAACAGGCCCAAGTCCCTGGGCTCCGAAGCAGAAAGCGCTGAACAGTAAATAG	780
QY	781	GCACAGAGAGAGGCGCAGTTCGATTTCTGTGGGAAGGCCAAGGCTAAAGCTGTAATTCGA	840
Db	781	GCACAGAGAGAGGCGCAGTTCGATTTCTGTGGGAAGGCCAAGGCTAAAGCTGTAATTCGA	840
QY	841	ATCTGCGCTGACAGCTCTGACACACAATAATGAGATGACGACGCTTCACTACTGTGGCC	900
Db	841	ATCTGCGCTGACAGCTCTGACACACAATAATGAGATGACGACGCTTCACTACTGTGGCC	900
QY	901	GAGCAGTATGTCAAGCGGTTCTCAAAACGTGGGCAAGGATTCACACAGTATCCTACTGGCC	960
Db	901	GAGCAGTATGTCAAGCGGTTCTCAAAACGTGGGCAAGGATTCACACAGTATCCTACTGGCC	960
QY	961	TCCAAACCTTGCGATGTCAACAGCATGTTGGGCTCAGGGCCATGGGTATATGAGACCCTC	1020
Db	961	TCCAAACCTTGCGATGTCAACAGCATGTTGGGCTCAGGGCCATGGGTATATGAGACCCTC	1020
QY	1021	ACCAAAACCCCAAGTGCAGAGGATCCAGACTACTCTTCCAGTGTGGAGCAGCAGAGATGTC	1080
Db	1021	ACCAAAACCCCAAGTGCAGAGGATCCAGACTACTCTTCCAGTGTGGAGCAGCAGAGATGTC	1080

0y	1081	CAGGGTCACAGTCGAACCTNTTGATGTGGAAACCTTGATCGGTACAAATAGTAGTATGAGGAC	1140
	1081	CAGGGTCACAGTCGAACCTNTTGATGTGGAAACCTTGATCGGTACAAATAGTAGTATGAGGAC	
Db	1081	CAGGGTCACAGTCGAACCTNTTGATGTGGAAACCTTGATCGGTACAAATAGTAGTATGAGGAC	1140
0y	1141	TGGGCTTNGCCAGGAGACTCTGGGGACACAGGACAGATTTTCTCATTT	1188
	1141	TGGGCTTNGCCAGGAGACTCTGGGGACACAGGACAGATTTTCTCATTT	
Db	1141	TGGGCTTNGCCAGGAGACTCTGGGGACACAGGACAGATTTTCTCATTT	1188

RESULT 2
US-08-232-463-14/c

; Sequence 14, Application US/08232463
; Patent NO. 5670367

```

; GENERAL INFORMATION:
;
; APPLICANT: DORNER, F.

```

```

; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
;

```

```

; TITLE OF INVENTION: RECOMBINANT FOLYPOX
;
; NUMBER OF SEQUENCES: 52
;

```

;; CORRESPONDENCE ADDRESS;
;; ADDRESSEE: Foley & Lardner
;;
***** 1000 Delaware Street, Suite 1400

CITY: Alexandria

COUNTRY: USA
ZTR: 32313-0280

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Vers
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/232,463
FILING DATE:

```

; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:

```

;; APPLICATION NUMBER: US/07/935,313
;; FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 30 700

REGISTRATION NUMBER: 25,108
REFERENCE/DOCKET NUMBER: 30472/114
TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109

TELEPHONE: (703) 200-7100
TELEX: 899149
INFORMATION FOR SEO ID NO: 14:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
;

```

```

;      TOPOLOGY: linear
;
IMMEDIATE SOURCE:
;

```

CLONE: PT29pt-F1s
US-08-232-463-14

Query Match	4.1%; Score 48.8;
Best Local Similarity	2.6%; Pred. No. 0.

Matches 8; Conservative 184; Mismatch

```
qy 01/ GGGGAAAGAGICIAIGCAGAIUCAGGUGGAGGCCA
:: :::::: : : :: : :: : :::: :
pb 1363 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
```

537 TAGACTTCAGCGGACCGAGACTGCGCGATCAAT

[illegible]

737 AGATCCTGGCCTCCGAGCAGAAAAGGCTGAACAG


```

RESULT 5
US-08-221-817-12
; Sequence 12, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoeftstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; TITLE OF INVENTION: kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532151and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2204 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1758
; US-08-221-817-12

Query Match 3.3%; Score 39; DB 1; Length 2204;
Best Local Similarity 60.6%; Pred.No.0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0

QY 557 CTGACGCTGGGGTATCCCTGCGCTTCGTCATGATCAAGATATTCATGTCACACCC 616
DB 1127 CTGACTGGTGGGGCGCTGGCTGCTCTCTGACAGATGATCGAGGCGACGTCCCTTCC 1186
QY 617 GGGTCAAAAGACTTATGACAGATCAGAGTGGAGGCGAGCGCGG 660
DB 1187 AGCAGAGGAAGAAGATCAACGCGGAGAGAGGTGAGCGGCTG 1230

RESULT 6
US-08-454-439-12
; Sequence 12, Application US/08454439
; Patent No. 5501618
; GENERAL INFORMATION:

```

```

APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoekstra, Merle F.
TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5591618and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2204 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1758
US-08-454-439-12

Query Match 3.3%, Score 39; DB 1; Length 2204;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0.

QY 557 CTGACTGCTGGGGGTCGCTGCCTGCTATGATGATCAAGATATCCATGTCACACCC 616
      ||||| |||| | ||||| | | |||| | | |||| | | |||| | | |||| |
Db 1127 CTGACTGGGGGCTCGCTGCTCTCTCTCTACGAGATGATCCAGGACGTCGCCCTTCC 1186
      ||||| |||| | ||||| | | |||| | | |||| | | |||| | | |||| |

QY 617 GGGTGAAGAGCTCTATGCAGATGCAGGTGAGGACGAGCGGCGG 660
      || || | | | | | | | | | | | | | | | | | | | |
Db 1187 AGCAGAGGAAGAAGATCAAGCGGAGAGAGGTGAGCGGCTG 1230

RESULT 7
PCT-US94-10487-12
; Sequence 12, Application PC/JUS9410487
; GENERAL INFORMATION:
; APPLICANT: ICOS Corporation
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;

```

```

: ADDRESSEE: Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/10487
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/221,817
: FILING DATE: 31 MAR 1994
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/123,932
: FILING DATE: 17 SEP 1993
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Noland, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 27866/31981
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2204 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31..1758
: PCT-US94-10487-12
:
: Query Match 3.3%; Score 39; DB 5; Length 2204;
: Best Local Similarity 60.6%; Pred. No. 0.17;
: Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
:
: QY 557 CTGACTGCTGGGGTATCCGCTCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCC 616
: ||||| |||| | ||||| | ||||| | ||||| | ||||| |
: DB 1127 CTGACTGGTGGCGCTCGCGCTCCTCTGTACGAGATGATCGACGCCAGTGGCCCTTCC 1186
:
: QY 617 GGGTAAAGAGCTCTATGCAGATGCAGGTGAGGCGAGCGCGCG 660
: || || || || || || || || || || || || || || || || ||
: DB 1187 AGCAGAGGAAGAAGATCAAGCGGAGAGAGTGTGAGCGGCTG 1230
:
: RESULT 8
: US-08-221-817-10
: Sequence 10, Application US/08221817
: Patent No. 5532151
: GENERAL INFORMATION:
: APPLICANT: Chantry, David
: APPLICANT: Gray, Patrick W.
: APPLICANT: Hoekstra, Merle F.
: TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSER: Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago

```

```

: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/221,817
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/123,932
: FILING DATE: 17 SEP 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5532151and, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 31981
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2206 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31..1926
: US-08-221-817-10
:
: Query Match 3.3%; Score 39; DB 1; Length 2206;
: Best Local Similarity 60.6%; Pred. No. 0.17;
: Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
:
: QY 557 CTGACTGCTGGGGTATCCGCTCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCC 616
: ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |
: DB 1124 CTGACTGGTGGCGCTCGCGCTCCTCTGTACGAGATGATCGACGCCAGTGGCCCTTCC 1183
:
: QY 617 GGGTAAAGAGCTCTATGCAGATGCAGGTGAGGCGAGCGCGCG 660
: || || || || || || || || || || || || || || || || ||
: DB 1184 AGCAGAGGAAGAAGATCAAGCGGAGAGAGTGTGAGCGGCTG 1227
:
: RESULT 9
: US-08-454-439-10
: Sequence 10, Application US/08454439
: Patent No. 5591618
: GENERAL INFORMATION:
: APPLICANT: Chantry, David
: APPLICANT: Gray, Patrick W.
: APPLICANT: Hoekstra, Merle F.
: TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSER: Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5591618and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO. 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2206 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1926
US-08-454-439-10

```

```

Query Match Similarity      3.3%; Score 39; DB 1; Length 2206;
Best Local Similarity      60.6%; Pred. No. 0.17;
Matches      63; Conservative 0; Mismatches 41; Indels 0; Gaps 0.

OY      557      CTGACTGCTGGGGTATCCGCTGCCTNCGTTATGAGATCAAGATATCCATGTGCCACCCC      616
          ||||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1124      CTGATCGTGCGGCCCTCGCGCTCTCCCTGTAGCAATGATGCCAGGCCAGTGCCTTCC      1183
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY      617      GGGTGAAAGAGCTCTTGCAGTGCAGTGCAGTGCAGGAGGAGCGGGG      660
          | || | || | || | || | || | || | || | || | || | || | || |
Db      1184      AGCAGAGCAAGACAAGATCAAGCCGGAGAGGTGGAGCGGCTG      1227

RESULT      10
PCT-US94-10487-10
: Sequence 10, Application PC/TUS9410487
: GENERAL INFORMATION:
: APPLICANT: ICOS Corporation
: TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
: TITLE OF INVENTION: Kinase GRK6
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/10487
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/221,817
: FILING DATE: 31 MAR 1994
: CLASSIFICATION:

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/123,932
 FILING DATE: 17 SEP 1993
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Noland, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/31981
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEF: 25-3856
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2206 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 31..1926
 PCT-US94-10487-10

```

Query Match Similarity      3.3%: Score 39; DB 5; Length 2206;
Match Local Similarity     60.6%: Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 557 CTGACTGCTGGGGTATCCGCTGCCCTNCGTTATAGATCAAGATATCCATGTGCCACCCC 616
      ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |
DB 1124 CTGATGCTGGCGCGCTCGCTCTCTCTCTGTACGATGATCGCAGGCCAGTGCCTTCC 1183
      ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |

OY 617 GGGTAAAGACTCTATGCAGATGCAGGTGGAGGCGAGCGGGCG 660
      | || | || | || | || | || | || | ||||| ||||| |
DB 1184 AGCAAGAGGAAGAAAGATCAAGCGGGAGAGGTGAGCGGCTG 1227

RESULT 11
US-08-149-097D-23
; Sequence 23, Application US/08149097D
; Patent No. 5874236
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149, 097D
; FILING DATE: 05-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105, 536
; FILING DATE: 11-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US92/06903

```

```

: FILING DATE: 14-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/914,231
: FILING DATE: 13-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/868,354
: FILING DATE: 10-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-55038
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 238-0999
: TELEFAX: (619) 238-0062
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7791 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 240..7037
: OTHER INFORMATION: /product="Alpha1A-2 subunit of
: OTHER INFORMATION: human calcium channel"
US-08-149-097D-23

Query Match          3.3%; Score 39; DB 2; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31;
Matches 111; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 564 CTGGGGTATCCCGCTGCTTCGTTATGATGATCAAGATATCATCCAGTCCACCCGGGTGAA 623
DB 2333 CTTTGGGACTACACCCCTCTTAATGTGTTCTTGGCCATGCTGTGGACAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAAGATGAGGTGAGGCGAGAGCGCGGAACGGGCCACAGTTCTAGAGTC 683
DB 2393 CGCCAGAGAGCTCACCAAGGTGAGGCGGAGCAGCAGAAGGAAGAAGACGACCGAACA 2452
QY 684 TGAGGGAGCCGACAGAGTGGCCATCAATGTGGCAGAAAGGAAACAGGCCCGAGATCCT 743
DB 2453 GAAACTTGGCTTACAGAAAGCCAAAGAGGTGGCAGAAAGTGAATCTCTCTCCGGGCCAA 2512
QY 744 GGCCTCCGAGAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGAGCC 795
DB 2513 CATGCTATAGCTGTGAAGAGCAACAGAAATCAAAAAGCCAGCCAGACTCC 2564

RESULT 12
: Sequence 23, Application US/08949386
: Patent No. 6090623
: GENERAL INFORMATION:
```

```

: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: McCue, Ann
: APPLICANT: Gillespie, Alison
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: NUMBER OF INVENTIONS: METHODS
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: US
: ZIP: 92101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/949,386
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/290,012
: FILING DATE: 11-AUG-1994
: APPLICATION NUMBER: 08/149,097
: FILING DATE: 5-NOV-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/105,536
: FILING DATE: 11-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 519808
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 238-0999
: TELEFAX: (619) 238-0062
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7791 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 237..7037
: OTHER INFORMATION: /standard_name="Alpha-1A-2"
US-08-949-386-23

Query Match          3.3%; Score 39; DB 3; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31;
Matches 111; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 564 CTGGGGTATCCCGCTGCTTCGTTATGATGATCAAGATATCATGTCACCCGGGTGAA 623
DB 2333 CTTTGGGAACTACACCCCTCTTAATGTGTTCTTGGCCATGCTGTGGACAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAAGATGAGGTGAGGCGAGAGCGCGGAACGGGCCACAGTTCTAGAGTC 683
DB 2393 CGCCAGAGAGCTCACCAAGGTGAGGCGGAGCAGCAGAGAGAAAGAAAGACGACCGAACA 2452
QY 684 TGAGGGAGCCGACAGAGTGGCCATCAATGTGGCAGAAAGGAAACAGGCCCGAGATCCT 743
DB 2453 GAAACTTGGCTTACAGAAAGCCAAAGAGGTGGCAGAAAGTGAATCTCTCTCCGGGCCAA 2512
QY 744 GGCCTCCGAGAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGAGCC 795
DB 2513 CATGCTATAGCTGTGAAGAGCAACAGAAATCAAAAAGCCAGCCAGACTCC 2564
```

RESULT 13
US-08-450-562-23
Sequence 23, Application US/08450562
Patent No. 6096514
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Allison
APPLICANT: Feldman, Daniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,562
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,950
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,250
FILING DATE: 30-NOV-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/603,751
FILING DATE: 08-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 02-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 7791 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 237..7037
OTHER INFORMATION: //standard_name="Alpha-1A-2"
US-08-450-562-23
Query Match: 3.3%; Score 39; DB 3; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31; Mismatches 121; Indels 0; Gaps 0;
Matches 111; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 564 CTGGGGTATCCGCTCCCTGCTTATGATGATCAAGATATCCATGTCACCCGGGTGA 623
DB 2333 CTTTGGGAATCAGCCACCTGCTGATGTGTTGTCATCGCTGTGCAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAAGTGCAGGTGAGGCGGCGGGAAGCCACAGTTCTAGAGTC 683
DB 2393 CGCCCGAGAGCTCACCAAGGTGAGGCGGAGCAGCAAGAGAAGCAAGCAGCAACCA 2452
QY 684 TGAGGGAGCCGAGAGCTGCGCCATCATGTGCGCAGAGGAAGAAACAGCCAGATCCT 743
DB 2453 GAAACTTGCCCTACAGAAAGCCAGAGAGGTGCGAGAAGTGAAGTCTCTGTCTCGGCGCAA 2512
QY 744 GGCCTCCGAGCAGAAAGGCTGAACAGATTAATCAGGCAGCAGAGAGGCC 795
DB 2513 CATGCTATAGCTGTGAAGACACAGAAATCAAGAAAGCCAGCAAGTCC 2564
RESULT 14
US-08-984-709A-23
Sequence 23, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:

Db 2453 GAAAGTGGCCCTACAGAAAGCCAGAGAGTGCGACAGAAGTGAAGTCTCTGTGCGCGGCCAA 2512
Oy 744 GGCCCTCCGAGCAGAGAAAGGCTGAACAGATAAATCAGGCGAGCAGAGAGGCC 795
Db 2513 CATGCTATACCTGTGAAGAGAGCAACAGAGATCAAAAGCCAGCCCAAGTCC 2564

Search completed: September 22, 2002, 15:22:50
Job time: 6995 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:20:45 : Search time 2868.76 Seconds
(without alignments)
8959.362 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GGCTTCGGAGACNACCGCT.....GGAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US06005.COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06006.COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080.COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081.COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082.COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083.COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084.COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085.COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086.COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087.COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088.COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089.COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090.COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091.COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092.COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093.COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094.COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095.COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B.COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C.COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D.COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A.COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B.COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C.COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D.COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E.COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A.COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B.COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C.COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A.COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B.COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C.COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A.COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B.COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C.COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US100.COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US101.COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US6000.COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US6001.COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US6002.COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US6003.COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US6004.COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US6005.COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6006.COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6007.COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6008.COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6009.COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6010.COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6011.COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6012.COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6013.COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6014.COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6015.COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6016.COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6017.COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6018.COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6019.COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6020.COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6021.COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6022.COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6023.COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6024.COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6025.COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6026.COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6027.COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6028.COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6029.COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6030.COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6031.COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6032.COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6033.COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6034.COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6035.COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6036.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1166	98.1	1322	1 PCT-US98-12125-102	Sequence 102, App
2	1166	98.1	1322	1 US-09-209-4628B-106	Sequence 106, App
3	1160	97.6	1398	1 PCT-US01-03800A-1001	Sequence 1001, Ap
4	1160	97.6	1398	1 PCT-US01-04098A-2212	Sequence 2212, Ap
5	1160	97.6	1398	1 US-09-471-275-865	Sequence 865, App
6	1160	97.6	1398	1 US-09-496-914A-8410	Sequence 8410, Ap
7	1160	97.6	1398	1 US-09-560-875A-8410	Sequence 8410, Ap
8	1155	97.2	1384	17 US-09-338-467-941	Sequence 941, App
9	1155	97.2	1384	25 US-09-644-869-8024	Sequence 8024, Ap
10	1155	97.2	1384	25 US-09-649-162-7309	Sequence 7309, Ap
11	1155	97.2	1384	25 US-09-652-109-7925	Sequence 7925, Ap
12	1155	97.2	1384	25 US-09-652-127-7358	Sequence 7358, Ap
13	1155	97.2	1384	25 US-09-652-816-7419	Sequence 7419, Ap
14	1155	97.2	1384	25 US-09-652-913-8809	Sequence 8809, Ap
15	1155	97.2	1384	25 US-09-652-914-7997	Sequence 7997, Ap
16	1155	97.2	1384	31 US-09-808-384-941	Sequence 941, App
17	1144.4	96.3	1337	1 PCT-US98-12125-45	Sequence 45, App
18	1144.4	96.3	1337	16 US-09-209-4628A-45	Sequence 45, App
19	1144	96.3	1416	1 PCT-US01-04098A-244	Sequence 244, App
20	1144	96.3	1416	22 US-09-598-075-249	Sequence 249, App
21	1144	96.3	1416	22 US-09-598-075A-249	Sequence 249, App
22	1105.4	93.0	1429	56 US-60-172-373-10483	Sequence 10483, A
23	1104.8	93.0	9098	1 PCT-US01-04942A-34	Sequence 34, App
24	1104.8	93.0	9098	22 US-09-596-193A-37	Sequence 37, App
25	1086.4	91.4	1234	55 US-60-164-285-3468	Sequence 3468, Ap
26	1072	90.2	1308	16 US-09-205-070-13379	Sequence 13379, A
27	1072	90.2	1308	17 US-09-340-633-13379	Sequence 13379, A
28	1072	90.2	1308	33 US-09-898-888-13379	Sequence 13379, A
29	1067.6	89.9	1501	16 US-09-898-888A-13379	Sequence 13379, A
30	1067.6	89.9	1501	16 US-09-298-733-35	Sequence 35, App
31	1067.6	89.9	1501	16 US-09-298-733A-35	Sequence 35, App


```

1 RESULT 2
2 US-09-209-462B-106
3 ; Sequence 106, Application US/09209462B
4 ; GENERAL INFORMATION:
5 ; APPLICANT: Rosen et al.
6 ; TITLE OF INVENTION: 86 Human Secreted Proteins
7 ; FILE REFERENCE: P2008P1
8 ; CURRENT APPLICATION NUMBER: US/09/209,462B
9 ; CURRENT FILING DATE: 1998-12-11
10 ; PRIOR APPLICATION NUMBER: PCT/US98/12125
11 ; PRIOR FILING DATE: 1998-06-11
12 ; PRIOR APPLICATION NUMBER: 60/049,547
13 ; PRIOR FILING DATE: 1997-06-13
14 ; PRIOR APPLICATION NUMBER: 60/049,548
15 ; PRIOR FILING DATE: 1997-06-13
16 ; PRIOR APPLICATION NUMBER: 60/049,549
17 ; PRIOR FILING DATE: 1997-06-13
18 ; PRIOR APPLICATION NUMBER: 60/049,550
19 ; PRIOR FILING DATE: 1997-06-13
20 ; PRIOR APPLICATION NUMBER: 60/049,556
21 ; PRIOR FILING DATE: 1997-06-13
22 ; PRIOR APPLICATION NUMBER: 60/049,606
23 ; PRIOR FILING DATE: 1997-06-13
24 ; PRIOR APPLICATION NUMBER: 60/049,607
25 ; PRIOR FILING DATE: 1997-06-13
26 ; PRIOR APPLICATION NUMBER: 60/049,608
27 ; PRIOR FILING DATE: 1997-06-13
28 ; PRIOR APPLICATION NUMBER: 60/049,609
29 ; PRIOR FILING DATE: 1997-06-13
30 ; PRIOR APPLICATION NUMBER: 60/049,610
31 ; PRIOR FILING DATE: 1997-06-13
32 ; PRIOR APPLICATION NUMBER: 60/049,611
33 ; PRIOR FILING DATE: 1997-06-13
34 ; PRIOR APPLICATION NUMBER: 60/050,901
35 ; PRIOR FILING DATE: 1997-06-13
36 ; PRIOR APPLICATION NUMBER: 60/052,989
37 ; PRIOR FILING DATE: 1997-06-13
38 ; PRIOR APPLICATION NUMBER: 60/051,919
39 ; PRIOR FILING DATE: 1997-07-08
40 ; PRIOR APPLICATION NUMBER: 60/055,984
41 ; PRIOR FILING DATE: 1997-08-18
42 ; PRIOR APPLICATION NUMBER: 60/058,665
43 ; PRIOR FILING DATE: 1997-09-12
44 ; PRIOR APPLICATION NUMBER: 60/058,668
45 ; PRIOR FILING DATE: 1997-09-12
46 ; PRIOR APPLICATION NUMBER: 60/058,669
47 ; PRIOR FILING DATE: 1997-09-12
48 ; PRIOR APPLICATION NUMBER: 60/058,750
49 ; PRIOR FILING DATE: 1997-09-12
50 ; PRIOR APPLICATION NUMBER: 60/058,971
51 ; PRIOR FILING DATE: 1997-09-12
52 ; PRIOR APPLICATION NUMBER: 60/058,972
53 ; PRIOR FILING DATE: 1997-09-12
54 ; PRIOR APPLICATION NUMBER: 60/058,975
55 ; PRIOR FILING DATE: 1997-09-12
56 ; PRIOR APPLICATION NUMBER: 60/060,834
57 ; PRIOR FILING DATE: 1997-10-02
58 ; PRIOR APPLICATION NUMBER: 60/060,841
59 ; PRIOR FILING DATE: 1997-10-02
60 ; PRIOR APPLICATION NUMBER: 60/060,844
61 ; PRIOR FILING DATE: 1997-10-02
62 ; PRIOR APPLICATION NUMBER: 60/060,865
63 ; PRIOR FILING DATE: 1997-10-02
64 ; PRIOR APPLICATION NUMBER: 60/061,059
65 ; PRIOR FILING DATE: 1997-10-02
66 ; PRIOR APPLICATION NUMBER: 60/061,060
67 ; PRIOR FILING DATE: 1997-10-02
68 ; NUMBER OF SEQ ID NOS: 737
69 ; SOFTWARE: PatentIn Ver. 2.0
70 ; SEQ ID NO 106
71 ; LENGTH: 1322
72 ; TYPE: DNA
73 ; ORGANISM: Homo sapiens

```

[illegible]

|||||
Db 780 agcagagagcagctgacgtctctgacgagcgaagcgaagctgagacgtatcgaat 839
Oy 843 CCTGGCTCAGCTCTGACACACATATAGAGATGACAGCTTGAACACTGCTGGCGA 902
Db 840 cctggctcagctcgaacacataatgagatgacagacttgaactcagctgagcga 899
Oy 903 GCAGATGTCACGCGCTTCTCCAACTGGCCCAAGACTCAACAGATCTACTGCTC 962
Db 900 gagatctgacagcgctctcgaactgacgaagctcgaagctcgaagctcgaagctc 959
Oy 963 CAACCTGGCGATGTCACAGATGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 960 caacctgagctgacacagcagctgagctgagctgagctgagctgagctgagct 1019
Oy 1023 CAAGCCCGCAGGACGAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Db 1020 caaacgccccagctgagcagcagcagcagcagcagcagcagcagcagcagcag 1079
Oy 1083 GCGTACAGATGCACTGTTGATGAGAGAACTTGTATGAGTGAAGTGAAGTGA 1142
Db 1080 ggcgtacagctgacagctgagctgagcagcagcagcagcagcagcagcagcag 1139
Oy 1143 GCGTTCGCGAGGAGCTGCTGGGACCAAGAGAGATTTTCTGAT 1188
Db 1140 ggcctggcagagagctgctgagcagcagcagcagcagcagcagcagcagcagc 1185

RESULT 3

PCT-US01-03800A-1001/c
; Sequence 1001, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1001
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-03800A-1001

Query Match 97.6%; Score 1160; DB 1; Length 1398;
Best Local Similarity 99.4%; Pred. No. 2.6e-278;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 1 GCGTTCGCGAGCACCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1393 GCGTTCGCGAGCAGCAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1334
Oy 61 GAAATGCTGGCG 120
Db 1333 GAAATGCTGGCG 1275
Oy 121 GCGTTCG 179
Db 1274 GCGTTCG 1215
Oy 180 GTTCGTCG 229
Db 1214 GTTCGTCG 1155
Oy 240 CGAGCGCTGTTGAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 299
|||||

Db 1154 GGAGCGCTGTTGAACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1095
Oy 300 CAAGGAAATTTGTCATCAACGCTGCTGAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCT 359
Db 1094 CAAGGAAATTTGTCATCAACGCTGCTGAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCT 1035
Oy 360 GCAAAATCGATGAGTCTCTTACCTGCGCATATGAGACCTTTACAAGGACGCTGCT 419
Db 1034 GCAAAATCGATGAGTCTCTTACCTGCGCATATGAGACCTTTACAAGGACGCTGCT 975
Oy 420 GGAGGACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 974 GGAGGACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
Oy 480 CAACCTCTCTGTCGCAAAAGTCTTCCGGGAAAGGAGTCTGCTGCTGCTGCTGCTGCT 539
Db 914 CAACCTCTCTGTCGCAAAAGTCTTCCGGGAAAGGAGTCTGCTGCTGCTGCTGCTGCT 855
Oy 540 TCCCATCAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
Db 854 TCCCATCAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Oy 600 TATCCATGTCGACCCCGGCTGAAAGAGTCTATGAGATGAGTGGAGGAGGAGCGGCG 659
Db 794 TATCCATGTCGACCCCGGCTGAAAGAGTCTATGAGATGAGTGGAGGAGGAGCGGCG 735
Oy 660 GAAAGGCGCCACAGTCTGATGAGTGTGAGGAGCCGAGAGTGGCCATCAATGTGGCAGA 719
Db 734 GAAAGGCGCCACAGTCTGATGAGTGTGAGGAGCCGAGAGTGGCCATCAATGTGGCAGA 675
Oy 720 AGGGAAGAAACGCGCCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
Db 674 AGGGAAGAAACGCGCCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
Oy 780 GCGAGCAGGAGAGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
Db 614 GCGAGCAGGAGAGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
Oy 840 AATCTGCTGAGCTGTGACACACATATATGAGATGAGCAGAGCTTCACTGCTGCTGCT 899
Db 554 AATCTGCTGAGCTGTGACACACATATATGAGATGAGCAGAGCTTCACTGCTGCTGCT 495
Oy 900 CGAGCAGTATGTCAGCGGCTTCTCCAACTGGCCAAAGGAGCTCAACATCTCTGCTGCT 959
Db 494 CGAGCAGTATGTCAGCGGCTTCTCCAACTGGCCAAAGGAGCTCAACATCTCTGCTGCT 435
Oy 960 CTCGAACCTTGGGATGTGACAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
Db 434 CTCGAACCTTGGGATGTGACAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375
Oy 1020 CACCAAGCCCGCAGTCCAGGAGCTCAGACTCAGCTCTCACTGGGAGCAGCAGATGCT 1079
Db 374 CACCAAGCCCGCAGTCCAGGAGCTCAGACTCAGCTCTCACTGGGAGCAGCAGATGCT 315
Oy 1080 CCAAGGTACAGATGCAAGTNTGATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
Db 314 CCAAGGTACAGATGCAAGTNTGATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255
Oy 1140 CTGGGCTTNGCAGGAGTCTGGGGAACAAGAGAGATTTTCTGAT 1188
Db 254 CTGGGCTTNGCAGGAGTCTGGGGAACAAGAGAGATTTTCTGAT 206
|||||

RESULT 4

PCT-US01-04098A-2212/c
; Sequence 2212, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned

```

? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: 09/728,422
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: 09/693,325
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 09/663,561
? PRIOR FILING DATE: 2000-09-15
? PRIOR APPLICATION NUMBER: 09/654,936
? PRIOR FILING DATE: 2000-09-01
? PRIOR APPLICATION NUMBER: 09/620,325
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: 09/598,075
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: 09/560,875
? PRIOR FILING DATE: 2000-04-27
? PRIOR APPLICATION NUMBER: 09/496,914
? PRIOR FILING DATE: 2000-02-03
? NUMBER OF SEQ ID NOS: 3960
? SOFTWARE: Custom
? SEQ ID NO 2212
? LENGTH: 1398
? TYPE: DNA
? ORGANISM: Homo sapiens
PCT-US01-04098A-2212
```

```

Query Match          97.6%; Score 1160; DB 1; Length 1398;
Best Local Similarity 99.4%; Pred. No. 2,6e-278;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```

OY 1 GGCCTTCGGAGCAGCCGCTCCGCTCGTCTGTTGGTCCGAGAGTCGTCGCGCGGTGG 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1393 GCGTTCGGAGCAGCCGCTCCGCTCGTCTGTTGGTCCGAGAGTCGTCGCGCGGTGG 1334
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1333 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1275
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 121 GGCCTTCGGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1274 GCGTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1215
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 180 GTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1214 GTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1155
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 240 GAGGCGCTGGTTGAACATCTCTCATCCCTGTTAGACCCGATCCGATATGTGAGAGTCT 299
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1154 GAGGCGCTGGTTGAACATCTCTCATCCCTGTTAGACCCGATCCGATATGTGAGAGTCT 1095
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 300 CAAGAAATTTGTCATCAACGTCGCTGAGCAGTGGCTGTGACTCTCGACAAATGTAAGTCT 359
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1094 CAAGAAATTTGTCATCAACGTCGCTGAGCAGTGGCTGTGACTCTCGACAAATGTAAGTCT 1035
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 360 GCAATTCGATGAGAGTCCTTAACTCGCGCATCATGACACCTTACAAAGCAAGCTACGGTGT 419
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1034 GCAATTCGATGAGAGTCCTTAACTCGCGCATCATGACACCTTACAAAGCAAGCTACGGTGT 975
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 420 GAGGAGCCCTGAGTATGCGGTGACCCAGCTAGTCAAAACACCATGATGATCAAGCTCGG 479
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 974 GAGGAGCCCTGAGTATGCGGTGACCCAGCTAGTCAAAACACCATGATGATCAAGCTCGG 915
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 480 CAAACTCTGTTGAGCAAGTCTTCCGGGAACGGAGTCCCTGAATGCCAGCTTGTGGA 539
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 914 CAAACTCTGTTGAGCAAGTCTTCCGGGAACGGAGTCCCTGAATGCCAGCTTGTGGA 855
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 540 TGGCATCAACCAAGCTGCTGACTGCTGGGGGTATCCGCTGCCCTNCGTTTATGAGATCAAGGA 599
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 854 TGGCATCAACCAAGCTGCTGACTGCTGGGGGTATCCGCTGCCCTNCGTTTATGAGATCAAGGA 795
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 600 TATTCATGTGCCACCCGGGTGAAAGAGTGTATGACAGATGACAGTGGAGGCAAGGGGCG 659
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 794 TATTCATGTGCCACCCGGGTGAAAGAGTGTATGACAGATGACAGTGGAGGCAAGGGGCG 735
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```

OY 660 GAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCCATCATGTGCGAGA 719
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 734 GAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCCATCATGTGCGAGA 675
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 720 AGGAGAAACAGGCCACAGTCTCTGCGCTCCGAGCAGAAAGGCTGACAGATAAATCA 779
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 674 AGGAGAAACAGGCCACAGTCTCTGCGCTCCGAGCAGAAAGGCTGACAGATAAATCA 615
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 780 GGGAGCAGAGAGGCCAGTCCAGTCTGCGGGAAGGCCCAAGCTAAAGCTAAGCTATTTCG 839
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 614 GGGAGCAGAGAGGCCAGTCCAGTCTGCGGGAAGGCCCAAGCTAAAGCTAAGCTATTTCG 555
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 840 AATCTGCTGACAGTCTTGACACACATAATGAGATGACAGAGCTTCACTGACTGCGC 899
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 554 AATCTGCTGACAGTCTTGACACACATAATGAGATGACAGAGCTTCACTGACTGCGC 495
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 900 CGAGCAGTATGTACAGCGCGTTCCTCAAACTGCGCCAGAGACTCCAAACATATCTACTGCC 959
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 494 CGAGCAGTATGTACAGCGCGTTCCTCAAACTGCGCCAGAGACTCCAAACATATCTACTGCC 435
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 960 CTCGAACCCCTGGGATGTCAACAGCATGATGCTCAGAGCCATGGTGTATATGAGCCCT 1019
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 434 CTCGAACCCCTGGGATGTCAACAGCATGATGCTCAGAGCCATGGTGTATATGAGCCCT 375
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 1020 CACCAAGCCCGCAGTGCAGAGGACTCCAGACTCACTCTCCAGTGGAGCAGAGAGATGT 1079
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 374 CACCAAGCCCGCAGTGCAGAGGACTCCAGACTCACTCTCCAGTGGAGCAGAGAGATGT 315
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 1080 CCAAGGTACAGATGCAAGTGTGATGAGAGACTTGTATGATGCAAGATGATTAGTGAG 1139
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 314 CCAAGGTACAGATGCAAGTGTGATGAGAGACTTGTATGATGCAAGATGATTAGTGAG 255
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 1140 CTGGGCTTNGCCAGAGTCTGGGAGCAAGCAAGCATTTCTCTGATT 1188
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 254 CTGGGCTTNGCCAGAGTCTGGGAGCAAGCAAGCATTTCTCTGATT 206
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 5
US-09-471-275-865/c
? Sequence 865, Application US/09471275
? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? TITLE OF INVENTION: Novel Contigs Obtained
? TITLE OF INVENTION: From Various Libraries
? FILE REFERENCE: 782
? CURRENT APPLICATION NUMBER: US/09/471,275
? EARLIER FILING DATE: 1999-12-23
? EARLIER APPLICATION NUMBER: US 09/235,076
? EARLIER FILING DATE: 1999-01-20
? EARLIER APPLICATION NUMBER: US 09/234,611
? EARLIER FILING DATE: 1999-01-22
? EARLIER APPLICATION NUMBER: US 09/240,371
? EARLIER FILING DATE: 1999-01-29
? EARLIER APPLICATION NUMBER: US 09/277,227
? EARLIER FILING DATE: 1999-03-25
? EARLIER APPLICATION NUMBER: US 09/271,490
? EARLIER FILING DATE: 1999-03-18
? EARLIER APPLICATION NUMBER: US 09/293,972
? EARLIER FILING DATE: 1999-04-15
? EARLIER APPLICATION NUMBER: US 09/274,861
? EARLIER FILING DATE: 1999-03-23
? EARLIER APPLICATION NUMBER: US 60/125,453
? EARLIER FILING DATE: 1999-03-19
? EARLIER APPLICATION NUMBER: US 60/126,605
? EARLIER FILING DATE: 1999-03-26
? EARLIER APPLICATION NUMBER: US 09/306,350
? EARLIER FILING DATE: 1999-05-07
? EARLIER APPLICATION NUMBER: US 09/399,720
? EARLIER FILING DATE: 1999-09-21
? EARLIER APPLICATION NUMBER: US 09/404,284
? EARLIER FILING DATE: 1999-09-21
? EARLIER APPLICATION NUMBER:
```



```

; PRIOR APPLICATION NUMBER: US 09/274,861
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/125,453
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/126,605
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: US 09/306,350
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/399,720
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: US 09/404,284
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: US 09/465,877
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 10410
; SOFTWARE: pc_ct_genes Version 1.02
; SEQ ID NO 8410
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1296)...(264)
; OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTxy 3.3f00, default parameters
US-09-496-914A-8410
```

```

Query Match          97.6%  Score 1160; DB 18; Length 1398;
Best Local Similarity 99.4%  Pred. No. 2,6e-278;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```

QY 1 GCGTTCTGGAGACNACCGCTCCGCTGCTCTGCTGTTGTCGGAGGTCGCTCGCGCGGTG 60
    |||
DB 1393 GCGTTCTGGAGAGCGACGCTCCGCTGCTCTGCTGTTGTCGGAGGTCGCTCGCGCGGTG 1334
    |||

QY 61 GAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    |||
DB 1333 GAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1275
    |||

QY 121 GCGTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
    |||
DB 1274 GCGTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1215
    |||

QY 180 GTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
    |||
DB 1214 GTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1155
    |||

QY 240 GGAGCTGTTGTAACATCCTCATCCTGTTAGACCGGATCGATATGTCAGAGTCT 299
    |||
DB 1154 GGAGCTGTTGTAACATCCTCATCCTGTTAGACCGGATCGATATGTCAGAGTCT 1095
    |||

QY 300 CAAGAAATTTGTCATCAACGTCGCTGAGAGTGGCTGTGACTCTCGACAAATTAATCT 359
    |||
DB 1094 CAAGAAATTTGTCATCAACGTCGCTGAGAGTGGCTGTGACTCTCGACAAATTAATCT 1035
    |||

QY 360 GCAAAATGATGAGTCTTACTGTCGATCATGAGACCCCTTAAGGCAAGCTACGCTGT 419
    |||
DB 1034 GCAAAATGATGAGTCTTACTGTCGATCATGAGACCCCTTAAGGCAAGCTACGCTGT 975
    |||

QY 420 GGAGGACCTGATGATGCTGACCGGATGATGATCAAAACACCATGATGATGATGATG 479
    |||
DB 974 GGAGGACCTGATGATGCTGACCGGATGATGATCAAAACACCATGATGATGATGATG 915
    |||

QY 480 CAAACTCTCTGTCGACAAAGTCTTCCGGGAAAGGAGTCCCTGAATGCCAGATTTGGA 539
    |||
DB 914 CAAACTCTCTGTCGACAAAGTCTTCCGGGAAAGGAGTCCCTGAATGCCAGATTTGGA 855
    |||

QY 540 TGGCATCAACCAAGCTGTCGATGCTGTTGGGATGATGCTGCTGCTGCTGCTGCTGCT 599
    |||
DB 854 TGGCATCAACCAAGCTGTCGATGCTGTTGGGATGATGCTGCTGCTGCTGCTGCTGCT 795
    |||
```

```

QY 600 TATCATGTGTCACACCCCGGGTGAAGAAGTCTATGATGATGATGATGATGATGATG 659
    |||
DB 794 TATCATGTGTCACACCCCGGGTGAAGAAGTCTATGATGATGATGATGATGATGATG 735
    |||

QY 660 GAAAGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCCGATCAATGTGCGAGA 719
    |||
DB 734 GAAAGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCCGATCAATGTGCGAGA 675
    |||

QY 720 AGGGAAGAAACAGCCCAATCTGTCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 779
    |||
DB 674 AGGGAAGAAACAGCCCAATCTGTCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 615
    |||

QY 780 GGCAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 839
    |||
DB 614 GGCAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 555
    |||

QY 840 AATCTGGCTGACGCTCTGACACACATTAATGAGATGACGAGCTTCACTGACTGTGC 899
    |||
DB 554 AATCTGGCTGACGCTCTGACACACATTAATGAGATGACGAGCTTCACTGACTGTGC 495
    |||

QY 900 CGAGCATATGTCAGGCGCTTCTCCAACTGGCCCAAGGAGGAGGAGGAGGAGGAGGAG 959
    |||
DB 494 CGAGCATATGTCAGGCGCTTCTCCAACTGGCCCAAGGAGGAGGAGGAGGAGGAGGAG 435
    |||

QY 960 CTCACACCTGGGAGATGTCACACAGATGAGGCTCAGAGGAGGAGGAGGAGGAGGAGG 1019
    |||
DB 434 CTCACACCTGGGAGATGTCACACAGATGAGGCTCAGAGGAGGAGGAGGAGGAGGAGG 375
    |||

QY 1020 CACCAAGCCCGGAGTCCAGAGGAGTCCAGACTCTCTCCAGTGGAGGAGGAGGAGATGT 1079
    |||
DB 374 CACCAAGCCCGGAGTCCAGAGGAGTCCAGACTCTCTCCAGTGGAGGAGGAGGAGATGT 315
    |||

QY 1080 CCGAGGTACAGATGCAAGTCTTATGATGAGGAGTCTTATGATGAGGAGTCTTATG 1139
    |||
DB 314 CCGAGGTACAGATGCAAGTCTTATGATGAGGAGTCTTATGATGAGGAGTCTTATG 255
    |||

QY 1140 CTGGGCTTCCAGGAGGAGTCTGGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1188
    |||
DB 254 CTGGGCTTCCAGGAGGAGTCTGGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 205
    |||
```

```

RESULT 7
US-09-560-875A-8410/c
; Sequence 8410, Application US/09560875A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Slinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Confligs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 787
; CURRENT APPLICATION NUMBER: US/09/560,875A
; CURRENT FILING DATE: 2000-04-27
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10410
; SOFTWARE: pc_ct_genes Version 1.02
; SEQ ID NO 8410
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1296)...(264)
; OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTxy 3.3f00, default parameters
US-09-560-875A-8410
```

```

Query Match          97.6%  Score 1160; DB 22; Length 1398;
Best Local Similarity 99.4%  Pred. No. 2,6e-278;
```


Matches 1182: Conservative 0: Mismatches 5: Indels 2: Gaps 2:

```
OY 1 GGCCTCTGGAGCAGACCCGCTCCGCTGCTCGTGGTTCGAGAGCGCGTGGCGGCTG 60
DB 1393 GGCCTCTGGAGCAGACCCGCTCCGCTGCTCGTGGTTCGAGAGCGCGTGGCGGCTG 1334
OY 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 1333 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1275
OY 121 GGCCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 1274 GGCCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1215
OY 180 GTTCGTCGCGCGAGCAGAGAGCGCGTGGTGGAGCAGATGGCGCGATTCACCGATCT 239
DB 1214 GTTCGTCGCGCGAGCAGAGAGCGCGTGGTGGAGCAGATGGCGCGATTCACCGATCT 1155
OY 240 GGAGCCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGTTATGTGAGAGTCT 299
DB 1154 GGAGCCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGTTATGTGAGAGTCT 1095
OY 300 CAGGAAATGTGATCAGAGTGCCTGAGAGTGCCTGAGCTCTGAGCAATGTAACTCT 359
DB 1094 CAGGAAATGTGATCAGAGTGCCTGAGAGTGCCTGAGCTCTGAGCAATGTAACTCT 1035
OY 360 GCAATGATGAGAGTGCCTGAGAGTGCCTGAGAGTGCCTGAGAGTGCCTGAGAGTGC 419
DB 1034 GCAATGATGAGAGTGCCTGAGAGTGCCTGAGAGTGCCTGAGAGTGCCTGAGAGTGC 975
OY 420 GGAGGACCTGAGATGCGCGTACCAGCTAGCTCAACAGCATGAGATGAGAGTGC 479
DB 974 GGAGGACCTGAGATGCGCGTACCAGCTAGCTCAACAGCATGAGATGAGAGTGC 915
OY 480 CAAATCTCTGTTGAGCAAGTCTTCCGGAAGCGGAGTCCCTGAGAGCGCAGATTTGGA 539
DB 914 CAAATCTCTGTTGAGCAAGTCTTCCGGAAGCGGAGTCCCTGAGAGCGCAGATTTGGA 855
OY 540 TGCATCAACAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 599
DB 854 TGCATCAACAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 795
OY 600 TATCATGTGCAACCCCGGCGTGAAGAGTCTATGAGATGAGAGTGCAGAGCGCG 659
DB 794 TATCATGTGCAACCCCGGCGTGAAGAGTCTATGAGATGAGAGTGCAGAGCGCG 735
OY 660 GAAAGGGGCGACAGTCTGAGAGTGTGAGAGGAGCGGAGAGTGCAGATTTGAGCA 719
DB 734 GAAAGGGGCGACAGTCTGAGAGTGTGAGAGGAGCGGAGAGTGCAGATTTGAGCA 675
OY 720 AGGGAAGAAACAGGCGCAGATCTGCGCTCCGAGCAGAGAAAGGCGTGAAGCTATTTC 779
DB 674 AGGGAAGAAACAGGCGCAGATCTGCGCTCCGAGCAGAGAAAGGCGTGAAGCTATTTC 615
OY 780 GGCAGCAGAGAGGCGCAGTGTGAGTGTGCGGAGAGCGCAAGGCTAAAGCTATTTC 839
DB 614 GGCAGCAGAGAGGCGCAGTGTGAGTGTGCGGAGAGCGCAAGGCTAAAGCTATTTC 555
OY 840 AATTCCTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 899
DB 554 AATTCCTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 495
OY 900 CGAGCAGTATGTGAGGCGCTTCTCCAACTGGCGAAGAGCTCCAACTATCTACTGCG 959
DB 494 CGAGCAGTATGTGAGGCGCTTCTCCAACTGGCGAAGAGCTCCAACTATCTACTGCG 435
OY 960 CTCGCAACCTGGCGATGTGACAGAGTGTGCGTCAAGGCGATGGGTATATGAGAGCCCT 1019
DB 434 CTCGCAACCTGGCGATGTGACAGAGTGTGCGTCAAGGCGATGGGTATATGAGAGCCCT 375
OY 1020 CACCAAAAGCCGAGTGTGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1079
DB 374 CACCAAAAGCCGAGTGTGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 315
```

```
OY 1080 CCAGGCTACAGATGCAAGTNTTATGAGAGAACTGTATCGAGTCAAGATGATTAGTGAG 1139
DB 314 CCAGGCTACAGATGCAAGTNTTATGAGAGAACTGTATCGAGTCAAGATGATTAGTGAG 255
OY 1140 CTGGGCTTGGCCAGGAGTGTCTGGGAGCAGAGAGAGAGATTTCTGATTT 1188
DB 254 CTGGGCTTGGCCAGGAGTGTCTGGGAGCAGAGAGAGATTTCTGATTT 206
```

RESULT 8

```
US-09-338-467-941
; Sequence 941, Application US/09338467
; GENERAL INFORMATION:
; APPLICANT: Gearling, David P.
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-25P
; CURRENT FILING DATE: 1999-06-22
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 1022
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 941
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-338-467-941
```

Query Match 97.2% Score 1155; DB 17; Length 1384;

Best Local Similarity 99.4% Pred No. 4.5e-277;

Matches 1177: Conservative 0: Mismatches 5: Indels 2: Gaps 2:

```
OY 6 CTGGAGCAGACCCGCTCCGCTGCTGTTGTTCCGAGAGTGCCTGCGCGTGGGAAT 65
DB 4 CTGGAGCAGACCCGCTCCGCTGCTGTTGTTCCGAGAGTGCCTGCGCGTGGGAAT 63
OY 66 GGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
DB 64 GGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
OY 126 CTGGGCGCGCTCCG-CCGCGCGCTCTGATGATGCGCGGAAACACCGTGGTACTTTC 184
DB 123 CTGGGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
OY 185 TGCAGCAGAGAGGCGCTGAGTGTGAGAGCGAGTGGCGGATTCACCGGATCTTGAGC 244
DB 183 TGCAGCAGAGAGGCGCTGAGTGTGAGAGCGAGTGGCGGATTCACCGGATCTTGAGC 242
OY 245 CTGGTTGAAACATCCATCCCTGTTGTTGAGCGGATTCGATATGTGAGATCTCAAG 304
DB 243 CTGGTTGAAACATCCATCCCTGTTGTTGAGCGGATTCGATATGTGAGATCTCAAG 302
OY 305 AATTTGATCAACAGCGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 364
DB 303 AATTTGATCAACAGCGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 362
OY 365 TCGATGAGTCTCTTACATGCGCATCATGAGCCTTACAGGCAAGCTACGTTGTGAGAG 424
DB 363 TCGATGAGTCTCTTACATGCGCATCATGAGCCTTACAGGCAAGCTACGTTGTGAGAG 422
OY 425 ACCCTGAGTATGCGCTGACCCAGCTAGCTCAACAAACATGAGATGAGAGCTGGCAAC 484
DB 423 ACCCTGAGTATGCGCTGACCCAGCTAGCTCAACAAACATGAGATGAGAGCTGGCAAC 482
OY 485 TCTCTTGGACAAAGCTCTCCGGAACGAGAGTCCCTGTAATGCGACATGTGAGATGCA 544
DB 483 TCTCTTGGACAAAGCTCTCCGGAACGAGAGTCCCTGTAATGCGACATGTGAGATGCA 542
OY 545 TCAACCAAGCTGCTGACTGCTGGGATATCGCTGCTTCGTTATGAGATCAAGATATCC 604
```


|||||
Db 64 gctgcgcgcgcgcgcgc-99ggcactggggcccttctgtcgaaggcctctcctactgctt 122
126 TCGGCGCGGCTCGG-GCCGGGCTCTCTGTGATTTGCCCGCAACACCGTGTACTGTCG 184
123 ctggccgcgcctccgcgcgcgcctcctctcgtgattgccccgaacacgcgtgtacttctg 182
185 TGGCCGACGAGGAGGCGCTGGGTGGTGGAGCGAATGGGCGCATTCACCGGATCCTGAGC 244
183 tgcgcgacgaaggagccctgggtgtgtgtgaggaatggccgatccacccggatctctggagc 242
245 CTGGTTTGAACATCTCATCTCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAG 304
243 ctggtttgaacatccatccatccctctgtttagaccggatccgatatgtgcagagctcacaag 302
305 AATTGTTCATCAACGTCGCGTGGAGCAGTCGGCTGTGACTCTCGACATGTGTACTCTGAAA 364
303 aaattgtcatcaacgctgcgcgagcagtcgcgtgtgaactctcgcacaatgtatcctcaca 362
365 TCGATGGAGTCTTTTACCTGTGCGCATGTGAGACCCCTTACAAGGCAAGCTACGCTGGAGG 424
363 tgcgtaggtctcttacccttcacgcgcatatgaccccttaacgaagctaagctgctggaggt 422
425 ACCCTGAGTATGCGCTGACCCAGCTAGCTCAAAACAACATGAGATCAGAGCTCGCAAC 484
423 accctgagtagtcgcgtcacccagctagctcaaaacaacatgagatcagagctcgcgaac 482
485 TCGCTMTGACAAAGTCTTCCGGGAGCGGAGTCCCTGAATGCCACACTTMTGATGCCA 544
483 tccctcttgacaagaagctctccggaaaggagtccttgaaatgccaagatgtgtgattgcca 542
545 TCACCAAGTGTGCTGACTGCTGGGATATCCGCTGCTNCCTTATGATGATCAAGATATCC 604
543 tcaaccaagctgtgtcgtcgtcgtgtatccgcctcgcctcgttatgtatgaataagatacc 602
605 ATGTGCCACCCCGGGTGAAGAAGTCTATGACATGACATGAGTGAAGCAGCGCGGAAC 664
603 atgtgcacaccccggtgaaagagctatctagatgcaagtgagagcagagcgcggaac 662
665 GGGCCACAGTTCTAGAGTCTGAGGAGCGGAGCCGAGTCCGATCATGTCGCAAGAGGA 724
663 gggccacagttctagagctcgaaggagccgagtcggcctacatgTCGTCGCAAGAGGA 722
725 AGAAGACGGCCAGATCTGGCTCTCCGAAGCAGAAAAAGCTGAACAGATTAATCAGGCA 784
723 agaaacagcccaagatctcgcgcctccgaagcagaagaagctgaaacagataatcagggc 782
785 CAGGAGAGGCGACGTGACATTTGCGCAAGGCCAAGGCTAAAGCTGAAGCTATTGAAATCC 844
783 caggagaagcccaagctgagttctgcaaggccaagctlaaagctgaagctatcgaatcc 842
845 TGGCTGACACTCTGACAAACAATATGAGATGCACAGTTTACATGACATGTCGCGGAGC 904
843 tggctgacactctgacaaacaataatgagatgacagcagatcttcaacgactcgtgtccga 902
905 AGTATGTCAAGCGGCTCTCAAACTGGCCAAAGGACTCAACACTATCTACTGCCCTCA 964
903 agtatgtcaagcggtcttcaaaactgycgaaggaatccaaactatcctctcgtccctcca 962
965 ACCCTGGCATGTCAACAGCATGTGGCTCAGGSCATGGGTATATGAGCCCTGACCA 1024
963 accctggcatgtcaacagcatgtgtgctcaagcatalggtataltggtgcccctcacca 1022
1025 AAGCCCATGGCAGGAGCTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCAGG 1084
1023 aagcccatggcagagagcaccagactcaactccagtgaggacagagatgttccagg 1082
1085 GTACAGATGCAAGTNTGTAGAGAACTTGATGAGTCAAGATGAGTATAGAGCTGG 1144
1083 gtacagatgcaagtntgtagagaaacttgatgagtcgaatgagatcttggagagctggg 1142
1145 CTTNGCCAGGAGTCTGGGGACAAGACAGATTTTCTGATT 1186
|||||

Db 1143 ctggcccaaggagctctggggacaagaagcagatttctcctgatt 1186
RESULT 12
US-09-652-127-7358
; Sequence 7358, Application US/09652127
; GENERAL INFORMATION:
; APPLICANT: Shvian, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1183-001
; CURRENT APPLICATION NUMBER: US/09/652,127
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,134
; NUMBER OF SEQ ID NOS: 10475
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 7358
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-127-7358
Query Match 97.2%; Score 1155; DB 25; Length 1384:
Best Local Similarity 99.4%; Pred. No. 4,5e-277;
Matches 1177; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
QY 6 CTGGAGACNACGCTCCGCTGTCGTGTTGCCGAGGTGCTGCGGGTGGGAAT 65
|||||
Db 4 ctggagacgcgcctccgcgtcgtcgtctgtgtccgaagctcgtgcgtgcgtggaaat 63
66 GCMGGCGCGCGCGCGCGGCGACATGGGCGCTTTTGTGAGAGGGCTCTCTACTGGCTT 125
64 gctgcgcgcgcgcgcgc-99ggcactggggcccttctgtcgaaggctcgtgcgtgcgtggaaat 122
126 CTGGCGCGGCTCGG-GCCGGGCTCTCTGTGATTTGCCCGCAACACCGTGTACTGTCG 184
123 ctggccgcgcctccgcgcgcgcctcctctcgtgattgccccgaacacgcgtgtacttctg 182
185 TGGCCGACGAGGAGGCGCTGGGTGGTGGAGCGAATGGGCGCATTCACCGGATCCTGAGC 244
183 tgcgcgacgaaggagccctgggtgtgtgtgaggaatggccgatccacccggatccctggagc 242
245 CTGGTTTGAACATCTCATCTCTGTGTAGACCGGATCCGATATGTGATGTCGCAAG 304
243 ctggtttgaacatccatccatccctctgtttagaccggatccgatatgtgcagagctcaca 302
305 AATTGTTCATCAACGTCGCGTGGAGCAGTCGGCTGTGACTCTCGACATGTGTACTCTGAAA 364
303 aaattgtcatcaacgctgcgcgagcagtcgcgtgtgaactctcgcacaatgtatcctcaca 362
365 TCGATGGAGTCTTTTACCTGTGCGCATGTGAGACCCCTTACAAGGCAAGCTACGCTGGAGG 424
363 tgcgtaggtctcttacccttcacgcgcatatgaccccttaacgaagctaagctgctggaggt 422
425 ACCCTGAGTATGCGCTGACCCAGCTAGCTCAAAACAACATGAGATCAGAGCTGGCAAC 484
423 accctgagtagtcgcgtcacccagctagctcaaaacaacatgagatcagagctcgaac 482
485 TCGCTMTGACAAAGTCTTCCGGGAGCGGAGTCCCTGAATGCCACACTTMTGATGCCA 544
483 tccctcttgacaagaagctctccggaaaggagtccttgaaatgccaagatgtgtgattgcca 542
545 TCACCAAGTGTGACTGCTGGGATATCCGCTGCTNCCTTATGAGATCAAGATATCC 604
543 tcaaccaagctgtgtcgtcgtgtgtgtatccgcgtcgtcgtttagatcaagataatcc 602
605 ATGTGCCACCCCGGGTGAAGAAGTCTATGACATGACATGAGTGAAGCAGCGCGGAAC 664
603 atgtgcacaccccggtgaaagagctatctagatgcaagtgagagcagagcgcggaac 662
665 GGGCCACAGTTCTAGAGTCTGAGGAGACCGGAGTCCGATCATGTCGCAAGAGGA 724
|||||

? APPLICANT: Faib, Dean R.
 ? TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ? TITLE OF INVENTION: THEREFOR
 ? FILE REFERENCE: 1600.1174-001
 ? CURRENT APPLICATION NUMBER: US/09/652,913
 ? CURRENT FILING DATE: 2000-08-31
 ? PRIOR APPLICATION NUMBER: 60/152,107
 ? PRIOR FILING DATE: 1999-08-31
 ? NUMBER OF SEQ ID NOS: 10833
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 8809
 ? LENGTH: 1384
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? US-09-652-913-8809

Query Match	97.28;	Score 1155;	DB 25;	Length 1384;
Best Local Similarity	99.48;	Pred. No. 4.5e-277;		
Matches 1177;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 2

[illegible]

QY	785	CAGNAGAGGCGCATGTGACGTTCTTGCCGCAAGGCCAAGGCTTAACCTGAACGTATTGCAATCC	844
Db	783	caggagagagccagatgcagttctctggcgaagaagccaaagctctaaagctctaaagctatctgcaatcc	842
QY	845	TGGCTGCAGGCTCTTACACCACTAATATGAGAGATGCACACTTCAGTGACTGTGGCCGAGC	904
Db	843	ctgctctcagcctctgacacacaataatgagagatgcaacagcttcactgcactgctgagcctgagc	902
QY	905	AGTATGTCAGCGCGCTTCTCCAAACTGGCCCAAGGACTCCAAACTATCTACTGCGCTTCA	964
Db	903	agtatgtcagtcagcgctctccaactctgccaaggaatcccaacaactatctcactgcccctca	962
QY	965	ACCTCGGCGATGTACCAAGCATGTGTGGCTCAGGCCATGGGCTTATATGAGACCCCTACCA	1024
Db	963	acctctgcgagatgtacccaagatctggtctcagtcgacatggtgtatataatgagccctcacca	1022
QY	1025	AAGCCCCAGTCCCGAGGACATCCAGACTCCACTCTCCAGTGGGAGCAGACAGATGCCAGG	1084
Db	1023	aagcccccagtgccagggaaatctccagactcactcccaagtggagacagagaatgtccaa	1082
QY	1085	GTACAGATGCAAGTNTTGATGAGAGAACTTGATCGAATCAAGATGTAGTATGAGAGCTGGG	1144
Db	1083	gtacagatgcaagctcttgatgaggaactctgactgagtcgaagtgtagtgagagctggtg	1142
QY	1145	CTTMCACGAGGAGTCTGGGACACAAGACAGACAGATTTTCTGATTT	1188
Db	1143	cttggccaggagctcttgagacaagaataacagatcttctctgatt	1186

```

RESULT 15
US-09-652-914-7997
; Sequence 7997, Application US/09652914
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MO
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1193-001
; CURRENT APPLICATION NUMBER: US/09/652,914
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,112
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9677
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7997
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-914-7997

```

[illegible]

OY	305	AAATTGTCATCAACGTCGGCCGACCATGTGGCGTGTGACTCTGCACAATGTAATCTGCGAAA	364
Db	303	aaattgtcattcaacgtccggtgagctgagcttgagcttcgcacaatgttaacttcgcgaaa	362
OY	365	TCGATGAGACTCCCTTTACCTCCGCGCATGATGGACCTTTACAAAGGCAAGCTACGGTGTGGAGG	424
Db	363	tcgattgagactccctttaccttacccttcgcatacgaaccttacaaggcaagctcagtgctggagg	422
OY	425	ACCCCTGAGTATGGCGTTCACCCAGCTAGCTCAAMCAACCTATGAGATCAGACGTCGGCAAC	484
Db	423	acccttgagtatggcgtttcacccagcttagcttcaamcaacctatgagatcagacgtcggcaaac	482
OY	485	TCTCTNTGGACAAGTCTTCCGGGAACGGGAGTCCCTGATCCACGATTTGTGATGCCA	544
Db	483	tctctntggacaagcttccgggaaccttccggagggagtccttgatccacgatttgatgcca	542
OY	545	TCAACCAAGTCTGCTGACTGCTGGGGTATCCGCTGCTNCGTTATAGATCAAGATATCC	604
Db	543	tcaaccaagcttgcctgactgctggggtatccgctgctncgttatagatcaagatattcc	602
OY	605	ATGAGCACCCCGGGTGAAGAAGTCTATGCAGTGCAGGTGGAGGACAGCGCGCGAANC	664
Db	603	atgagcaccccggggtgaagaagtctatgcagtgctgaggtggagggacagcgcggaanc	662
OY	665	GGGCCACAGTTCTTAGAGCTTGAGGGGACCCGAGAGTGGGCCATCAATGTGGCAGAAAGGA	724
Db	663	ggggcacagttcttagagcttgaggggacccgagagtcggcccatcaatgtggcagaagga	722
OY	725	AGAAACGAGGCGCAGATCTCTGGCTCCGAAAGCAAGAAAGGCTACACAGATTAATCATGGCAG	784
Db	723	agaaacgagcgccagatctccggctccgaaagcagaagaaggctgaacaagataaactcaggcag	782
OY	785	CAGGAGAGGCGCAGTGCAGTTCTGGCGAAGGCCAAGGCTTAAAGCTGAAGCTATTTCGAATCC	844
Db	783	caggagagcgccagtgcagttcttcggcgaaaggccaaagcttaagcttagctatttcgaatcc	842
OY	845	TGGCTGACGCTTGACACACATTAATGAGAGATGCGACACTTCAGTGACTGTGGCCGAGC	904
Db	843	tggctgagcttgacacacattaatgagagatgcgacacttcagtgaactgtggccgagc	902
OY	905	AGATTCTACACGGGTTCTCCAAACTGGCCAGAGACTCACAACACTATTCAGTCCCTCCCA	964
Db	903	agattctcacgggttctccaaactggccagagactcacaacactattcagttccctccca	962
OY	965	ACCTGCGCATGTCCACCAGCATGCTGGGCTCAGGCCATGGGCTATATGGAGCCCTCACCA	1024
Db	963	acctgctcatgtccaccagcatgctgggctcaggccatgggctatatggagccctcacca	1022
OY	1025	AAGCCGAGTCCGAGGACTCCAGACGTCACCTCTCCAGTGGGAGCGACGAGATGTCCAGG	1084
Db	1023	aagccgagtccgaggtccagacctcagctcctccagtgggagcgacgagatgtccagg	1082
OY	1085	GTACAGATGCAAGTNTTGATGAGGAACCTTGATGCACTCAAGATGAGTTAGTGGAGCTGGG	1144
Db	1083	gtacagatgcaggtnttgatgaggaaccttgatgcactcaagatgagttagtggagctggg	1142
OY	1145	CTTGGCCAGGAGTCTGGGGACAAGAAGCAAGATTTTCTGATT	1188
Db	1143	cttggccaggagcttggggacaagaagcaagattttctgatt	1186

Search completed: September 22, 2002, 16:14:40
Job time: 6835 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:49:28 : Search time 322.77 Seconds
(without alignments)
11809.333 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCTTCTGGAGACNACGCT.....GGAAGCAGATTTCGTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2138461 seqs, 1604250230 residues

Total number of hits satisfying chosen parameters: 4276922

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCF_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	99.6	1188	5	US-09-898-216-2
2	1166	98.1	1322	6	US-10-219-793-106
3	1156.4	97.3	1910	6	US-10-198-846-12975
4	1145.8	96.4	1842	5	US-09-785-276A-24973
5	1144.4	96.3	1337	6	US-10-219-793-45
6	1130.4	95.2	1244	5	US-09-629-469A-11009
7	1104.8	93.0	9098	7	US-10-119-428-37
8	1086.4	91.4	1234	5	US-09-705-256A-3468
9	1057.4	89.0	1309	5	US-09-919-002-11863
10	815.6	68.7	1150	6	US-10-198-846-13409
11	618.6	52.1	622	1	PCF-US01-43704-1161
12	546.4	46.0	550	6	US-10-214-403-2995
13	538	44.3	563	1	PCF-US01-43704-58
14	516.6	43.5	691	7	US-10-097-105-1428
15	516.4	43.5	567	5	US-09-629-469A-1950
16	502	42.3	504	1	PCF-US01-43704-841
17	454	38.2	454	7	US-10-146-502-193
18	446	37.5	448	5	US-09-920-300A-1075
19	446	37.5	439	7	US-10-099-926-1075
20	423.4	34.6	439	1	PCF-US02-25766-1798
21	411.4	34.6	445	5	US-09-918-995-5335
22	394.4	33.2	478	5	US-09-918-995-31316
23	368.2	30.1	740	7	US-10-097-105-596
24	357	30.1	357	6	US-10-214-403-2873
25	352.8	29.7	362	5	US-09-721-544-8507

26	299.4	25.2	301	1	PCF-US02-12378-26	Sequence 26, Appl
27	299.4	25.2	301	7	US-10-124-805-26	Sequence 26, Appl
28	294	24.7	902	6	US-10-198-846-6734	Sequence 6734, Ap
29	292	24.6	300	5	US-09-539-331D-9558	Sequence 9558, Ap
30	289.4	24.4	521	1	PCF-US02-25766-8693	Sequence 8693, Ap
31	287	24.2	300	5	US-09-539-331D-21791	Sequence 21791, A
32	282.8	23.8	334	5	US-09-539-331D-28244	Sequence 28244, A
33	282.2	23.8	430	5	US-09-785-276A-37723	Sequence 37723, A
34	279.2	23.5	1407	5	US-09-886-492-17868	Sequence 17868, A
35	262	22.1	265	5	US-09-539-331D-14079	Sequence 14079, A
36	255	21.5	255	5	US-09-539-331D-4825	Sequence 4825, Ap
37	254.4	21.4	282	5	US-09-540-210B-26126	Sequence 26126, Ap
38	243.2	20.5	263	5	US-09-539-806B-2445	Sequence 2445, Ap
39	238	20.0	253	5	US-09-539-331D-688	Sequence 688, App
40	227.8	19.2	242	5	US-09-539-331D-10091	Sequence 30091, A
41	221	18.6	221	5	US-09-539-331D-15714	Sequence 15714, A
42	220	18.5	1123	6	US-10-219-999-8142	Sequence 8142, Ap
43	219	18.4	390	5	US-09-785-276A-7770	Sequence 7770, Ap
44	216.8	18.2	235	5	US-09-539-331D-29396	Sequence 29396, A
45	211	17.8	211	5	US-09-540-210B-18465	Sequence 18465, A

ALIGNMENTS

RESULT 1
US-09-898-216-2
Sequence 2, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898, 216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1188 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-898-216-2

	Query Match	99.68%	Score 1183;	DB 57;	Length 1188;
	Best Local Similarity	100.0%;	Pred. No. 1,9e-289;		
	Matches 1188;	Conservative	0;	Mismatches	0; Indels
					Gaps 0;
OY	1 GGCCTTGAGGACNMCCGGTCGCGTGTCTGTTGGATTCCGGAAGNCGCTGCGCGCGTGG				60
Db	1 GCCTTCGTGGAGACNMCCCTCCGCTCCTCTCTTGTTCCGAAAGTCCGTGCGCGCGTGG				60
OY	61 GAATGCTGGGCGCGCGCGCGCGCGCGACTGSGGCCCTTTTGTCTGAGGGGCTCTTACT				120
Db	61 GAATGCTGGGCGCGCGCGCGCGCGCGCGACTGSGGCCCTTTTGTCTGAGGGGCTCTTACT				120
OY	121 GGCCTTGAGGCGCGCGCTCGCGCGCGCTCTCTGTGATTGCCCGAACAACCTGGTACTG				180
Db	121 GGCCTTGAGGCGCGCGCTCGCGCGCGCTCTCTGTGATTGCCCGAACAACCTGGTACTG				180
OY	181 TTGCGTGGCCACAGAGAGGCTGGGGGTGGAGCGAATGGGCGGATTTCCACGGGATCTG				240
Db	181 TTGCGTGGCCACAGAGAGGCTGGGGGTGGAGCGAATGGGCGGATTTCCACGGGATCTG				240
OY	241 GAGCCTGGTTTTGAACATCCTCATCCCTGTGTAGAACC GGATCGSATAITGCAAGACTTC				300
Db	241 GAGCCTGGTTTTGAACATCCTCATCCCTGTGTAGAACC GGATCGSATAITGCAAGACTTC				300
OY	301 AAGGAATTTGTATCAACGCTGCCCTGAGCAGTGGGCTGTGACTCTGACAAATGTACTTG				360
Db	301 AAGGAATTTGTATCAACGCTGCCCTGAGCAGTGGGCTGTGACTCTGACAAATGTACTTG				360
OY	361 CAAATCGATGGAGTCTTTAACMGCGGCATATGAGACCCTTACAAAGCAAGCTACGGTGTG				420
Db	361 CAAATCGATGGAGTCTTTAACMGCGGCATATGAGACCCTTACAAAGCAAGCTACGGTGTG				420
OY	421 GAGGACCCCTGAAATATGCCCTCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGC				480
Db	421 GAGGACCCCTGAAATATGCCCTCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGC				480
OY	481 AAAGCTCTNTGGACAAAGCTTTCGGGGAGCGGAGTCCCCTGAATGCCAGATTGTGGAT				540
Db	481 AAAGCTCTNTGGACAAAGCTTTCGGGGAGCGGAGTCCCCTGAATGCCAGATTGTGGAT				540
OY	541 GCCATCAACCAAGCTGCTACGTGCGGGGATATCCGCTGCTNGTATGATCAAGAGAT				600
Db	541 GCCATCAACCAAGCTGCTACGTGCGGGGATATCCGCTGCTNGTATGATCAAGAGAT				600
OY	601 ATTCATGTGTCACCCCGGGGTGAAGAAGTATATGACAGATCAGGTGGAGCAGAGCGCGG				660
Db	601 ATTCATGTGTCACCCCGGGGTGAAGAAGTATATGACAGATCAGGTGGAGCAGAGCGCGG				660
OY	661 AAACGGGGCCACAGTTCTTAGAGTCTGAGGGGAGCCCGAGACTCGGCCATCAATGTGGCAA				720
Db	661 AAACGGGGCCACAGTTCTTAGAGTCTGAGGGGAGCCCGAGAGTCCGCCATCAATGTGGCAA				720
OY	721 GGGAGAAACAGGCCCCAGATTCCTGGGCTCCGAAAGCAAAAAAGGCTGAAACAGATAATTCAG				780
Db	721 GGGAGAAACAGGCCCCAGATTCCTGGGCTCCGAAAGCAAAAAAGGCTGAAACAGATAATTCAG				780
OY	781 GCAGCAGAGAGGCGCAGTCAAGTCTTGCGGGAAGGCCAAGGCTTAAGCTGAAGCTATTTCGA				840
Db	781 GCAGCAGAGAGGCGCAGTCAAGTCTTGCGGGAAGGCCAAGGCTTAAGCTGAAGCTATTTCGA				840
OY	841 ATTCGTGGCTGACAGCTCTGACACAACTAATGAGATGACAGCAGCTTCACTGACTGTGGCC				900
Db	841 ATTCGTGGCTGACAGCTCTGACACAACTAATGAGATGACAGCAGCTTCACTGACTGTGGCC				900
OY	901 GAGCAGTATGTCAAGCGGCTTCTCCAACCTGGCCAAAGGACTTCCAAACTATCTCACTGCCC				960
Db	901 GAGCAGTATGTCAAGCGGCTTCTCCAACCTGGCCAAAGGACTTCCAAACTATCTCACTGCCC				960
OY	961 TCCAAACCTGGGAGATTCACACAGATGGTGGCTCAGGACCATGGGTATATGAGAGCCCTC				1020
Db	961 TCCAAACCTGGGAGATTCACACAGATGGTGGCTCAGGACCATGGGTATATGAGAGCCCTC				1020
OY	1021 ACCAAAGCCCCAGTGGCAGGAGACTCCAGACTCACTCTCCAGTGGAGACAGCAGATGTC				1080

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

1      PRIOR APPLICATION NUMBER: 60/060,844
2      PRIOR FILING DATE: 1997-10-02
3      PRIOR APPLICATION NUMBER: 60/060,865
4      PRIOR FILING DATE: 1997-10-02
5      PRIOR APPLICATION NUMBER: 60/061,059
6      PRIOR FILING DATE: 1997-10-02
7      PRIOR APPLICATION NUMBER: 60/061,060
8      PRIOR FILING DATE: 1997-10-02
9      NUMBER OF SEQ ID NOS: 737
10     SOFTWARE: PatentIn Ver. 2.0
11     SEQ ID NO 106
12     LENGTH: 1322
13     TYPE: DNA
14     ORGANISM: Homo sapiens
15     FEATURE:
16     NAME/KEY: SITE
17     LOCATION: (82)
18     OTHER INFORMATION: n equals a,t,g, or c
19     FEATURE:
20     NAME/KEY: SITE
21     LOCATION: (1320)
22     OTHER INFORMATION: n equals a,t,g, or c
23     FEATURE:
24     NAME/KEY: SITE
25     LOCATION: (1321)
26     OTHER INFORMATION: n equals a,t,g, or c
27     FEATURE:
28     NAME/KEY: SITE
29     LOCATION: (1322)
30     OTHER INFORMATION: n equals a,t,g, or c
31     US-10-219-793-106

```

Query Match	98.1%	Score 1166	DB 6	Length 1322
Best Local Similarity	99.2%	Pred. No. 3.9e-285		
Matches 1177	Conservative	1	Mismatches 7	Indels 1

Oy	- 3	CTCTGGGAGGACGACGCTCCGCTGTCCTGTTGGTTCCGAGAGTCCTGCGCGGTGGGA	62
Dp	1	ctctcg9ag9agcagacgcctccgcctcgtctcgttctglttccg9aggtctcg9ctg9ctg9	60
Oy	63	AATGCTGGCGCGCGCGCGGCGGCGGAGCACTGGGGGCCCTTTTGCTAGAGGGCTCTCTATTGG	122
Dp	61	aatgcctgscgcgcgcgcgcgcgc- gmggcactg9ggcccttctgc9ag9gctctctactg9	119
Oy	123	CTTCTGGCCGCGCTCCGGCGCGCGCCTCCTCTGTGATTGCCCGGAACACCTGTGACTGT	182
Dp	120	ctctcgcgcgcgcctccgscgcgcctcctcctcgtgatttgcgccgaacacscgtgtactgtc	179
Oy	183	CGTGGCGGACGACGAGCGCTTGGTGTGTGAGCAATATGGCGGATTCACCGACTCTGGA	242
Dp	180	cgttccgcgcagcag9ag9ccttgc9ctgtgc9ag9aactg9gcgcattccacscgcgtctgc9a	239
Oy	243	GCCCTGTTTGAACATCCCATCCTCGTGTGTTAGACCGGATCCGATATGTGCAGAGTCCAA	302
Dp	240	gcccgtglttgcgaacatcccatccctcgtgtgtagaccgcgatacgcgaatgtagagactc	299
Oy	303	GGAAATTTTCATCAACCGCTGAGCAGTGGCGGTGATCACTCTGCACAAATGAACTTGGCA	362
Dp	300	ggaaattgtcatcaacgctgcctgcgcagctgcgcgtgtactctgcagacaatgtaactgcga	359
Oy	363	AATCGATGAGACTCCTTTTACCTGACGCGCATCATGAGACCTTTACAGGCAAGCTACGGTGTGGA	422
Dp	360	aatcgatgagctccttcttaccctgcgcgcacatcgtgaccttacaaggccaagcttaaggttgcga	419
Oy	423	GGACCCCTAAGTATGCGCTGCACCCAGCTAAGCTCAAAACAACATGAGATCAGAGCTGGCAA	482
Dp	420	ggacccttgaatgatgcgcgtcaaccagctagcttcaacaacaactgagatcagctgcgcgcaa	479
Oy	483	ACTCTCTMTGGACAAGGCTTTCGCGGAACGGGAGTCCCTGAATGCCAGCATTTGTGATGC	542
Dp	480	actctctcttgcgaacaagctcttccgcggaacgcgagctccttgaatgacgcagcatgtgcgagtc	539

OY	543	CATACCAACAGCTGCTACGTCGTCGGGGTATCCGCTCCCTCCGTTATAGATCAACGAT	602
OY	540	catcaaccaaacctgctgaactcgtctgggtatccacgtctcccttctatgataccaagatat	599
OY	603	CCATGTCGACCCCGGGGTGAAGAAGTCTATGACGATGACAGGTGAGAGCGACGGCGGAA	662
OY	600	ccatgtgcacccccgggttgaaagatctatgcagatgcacggttgagggcagagccggcga	659
OY	663	ACGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTGGCCCATCATGTTGGCAGAA	722
OY	660	acgggcacacagttctatagatctgagggagaccgcgagagtcgagcatcaatgctgcaga	719
OY	723	GAAGAACACAGGCCCATGATCCTGGCCCTCCGACGACGAGAAAGCGTGAACGATTAATC	782
OY	720	gaagaataaagccccagatccctggcctccgaaagcagaataaagctcgacaataaalc	779
OY	783	AGCAGAGAGAGCCAGTCTACGTTCTGGCGAAGGCCAAGGCTTAAGCTGAAGCTATTC	842
OY	780	agcagagagagccagctgcagttctctgcgaaagcccaagcttaagctcgaagcttcga	839
OY	843	CCTGGCTGCAGCTTGTGACACAAATATGAGATGACGACGCTTCACTGACTGTGGCG	902
OY	840	ctgtgcctgcagctctgcacacaataatgtagatgcaagcagcttcaactgactgtg	899
OY	903	GCAGATGTCAGCCGCTTTCACAAACGGGCGCAAGACGCTCAACACATATCTACGCTCC	962
OY	900	gcagatgctcagcgctctcccaaatctgcgaagagctccaacatctactcagctcc	959
OY	963	CAACCTTGGCGATGTCACACGATGCTGCGTCAGGCCATGGGTGTATATGAGCCCTC	1022
OY	960	caaccttgcgatgtctcacagcagatggtgtgcagcattggtgtatatgtgagccctc	1019
OY	1023	CAAGCCCGAGTGGCAGGAGACTCCAGACTACTCTCCATGGGGAGCAGAGAGATGTCA	1088
OY	1020	caaaagccccagctgcgaaggagctccagatctcccaagtgaggagcaagatgctca	1079
OY	1083	GGGTAACAGATCAAGTMTTCATGAGGAACCTGATCGATGCAAGATGATGTTACTGG	1142
OY	1080	gggtacagatgcaagctcttgaatgaagaaactgtatgagtcgaagaatgagttagtgag	1139
OY	1143	GGCTTGGCCAGGAGTCTGGGGACCAAGGAAGCAGATTTTCTGTGATT	1188
OY	1140	ggcttggccagggagctctggtgggaagagcagatltccgatt	1185
RESULT	3		
US-10-198-846-12975			
/ Sequence 12975, Application US/10198846			
/ GENERAL INFORMATION:			
/ APPLICANT: Illlie, James			
/ APPLICANT: Xu, Yongyao			
/ APPLICANT: Wang, Youzhen			
/ TITLE OF INVENTION: STEINMANN, KATHLEEN			
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER			
/ FILE REFERENCE: MRI-049			
/ CURRENT APPLICATION NUMBER: US/10/198,846			
/ CURRENT FILING DATE: 2002-07-18			
/ PRIOR APPLICATION NUMBER: 60/306,220			
/ PRIOR FILING DATE: 2001-07-18			
/ NUMBER OF SEQ ID NOS: 14084			
/ SOFTWARE: FastSeq for Windows Version 4.0			
/ SEQ ID NO 12975			
/ LENGTH: 1910			
/ TYPE: DNA			
/ ORGANISM: Homo sapiens			
/ FEATURE:			
/ NAME/KEY: misc_feature			
/ LOCATION: 1, 2, 1909, 1910			
/ OTHER INFORMATION: n = A,T,C or G			
US-10-198-846-12975			

```

RESULT 3
US-10-198-846-12975
: Sequence 12975, Application US/10198846
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS,
: TITLE OF INVENTION: FOR IDENTIFICATION,
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12975
: LENGTH: 1910
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..2,1909,1910
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12975

```

```

? TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
? TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
? TITLE OF INVENTION: THERAPY OF BREAST CANCER
? FILE REFERENCE: MRI-049
? CURRENT APPLICATION NUMBER: US/10/198,846
? CURRENT FILING DATE: 2002-07-18
? PRIOR APPLICATION NUMBER: 66/306,220
? PRIOR FILING DATE: 2001-07-18
? NUMBER OF SEQ. ID NOS: 14084
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ. ID NO 12975
? LENGTH: 1910
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc-feature
? LOCATION: 1, 2, 1909, 1910
? OTHER INFORMATION: n = A,T,C or G
? US-10-198-846-12975

```

OTHER INFORMATION: n = A, T, C or G
US-10-198-846-12975

Query Match	97.38;	Score 1156.4;	DB 6;	Length 1910;
Best Local Similarity	-99.38;	Pred. NO. 1.1e-282;		
Matches 1179; Conservative	0;	Mismatches 6;	Indels 2;	Gaps 2.

[illegible]

Qy	1022	CCAAAGCCCAAGTCCCAAGGACATCCAGACATATTCACAGTGGAGACAGACAGATGTCC	1081
Db	1022	ccaaagcccccgttgcagtggaacctccagctccacgtcccgatggcgacagagatgtcc	1081
Qy	1082	AGGGTACGATGACAGANTTGATGAGAGACTTATCGATCAAGTACGATTAGTGGAGCT	1141
Db	1082	agggtaacagatgcgaagctcttgcgaatgaggaacttgcatacgaatgagttctgttgagact	1141
Qy	1142	GGGGTTTNGCCAGGAGATCTGTGGGACAAAGACAGATTTTCTCGATT	1188
Db	1142	gggctcttgccagagagctctgtggacaaagagagagatttctctgatt	1188

```

RESULT 4
US-09-785-276A-24973
; Sequence 24973, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007B
; CURRENT FILING DATE: US/09/785, 276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183, 319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189, 862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207, 454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211, 314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219, 007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255, 281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 24973
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1841, 1842
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-24973

```

[illegible]

Dh 718 tggagccctgggttgaacatccatccctgctgttagaccggatccgatatgtgcaagatc 777
Oy 299 TCAGGAATTTGTCATCAACGTGCTGAGCAGTGGCTGTGACTCTCGACAATGTAATC 358
Dh 778 tcaaggaaattgtcatcaacagtgccgagcagtcgctgtgactctccgacaatgttaacc 837
Oy 359 TCAGAAATGATGAGTGTCTTTTACCTTCGCATCATGAGACCCCTTACAAGGCAAGCTACGGTG 418
Dh 838 tgcgaatcgatgagtccttaccctgcgcatcatgacaccttaacaggaagctacggtg 897
Oy 419 TGGAGGACCTGATGATGCGTCACCCAGCTAGCTCAACACACATGAGATGAGACTCG 478
Dh 898 tggaggaagccctgagtagtccgctcaacccagctagctcaacaaacatgagatcagagctcg 957
Oy 479 GCAAACTCTCTTGTGCAAAAGTCTTCCGGGAACGGGAGTCCCTGAATGCAAGCATGTGTG 538
Dh 958 gcaaacctctcttgcaacaaagcttccgggaacggagctcccgatgtccagatctgtg 1017
Oy 539 ATGCCATCAACCAAGCTGCTGACTGCTGGGATCCGCTGCTTGCCTTATGAGATCAAG 598
Dh 1018 atgcacatcaacaaagctgtagctgtgggtatccgctgcctccgtcttagatcaag 1077
Oy 599 ATATCCATGTGACCCCGGGGAAAGTATGATGAGATGAGATGAGATGAGATGAGATGAG 658
Dh 1078 atatcattatgctgcaaccccggtgaaagagctcatgagatgagatgagatgagatgag 1137
Oy 659 GGAACGCGCCACAGTCTGAGTGTGAGGAGACCCGAGATGCGGCATCAATGTGCGAG 718
Dh 1138 ggaacggcgccacagctctagagctggaaggagcccgagagctgagcatalcaatgtgccc 1197
Oy 719 AAGGAGAAGAACAGCCAGATCTGCTGCTCCGAAAGCAAAAAAGCTGAACAGATTAATC 778
Dh 1198 aagggaagaaagagcccgatccctgctccgaagcagaaagctgagacagataalc 1257
Oy 779 AGGACGACGAGAGGCGCATGCTGCTGCGAAGGCCAAGGCTTAAGCTGAAGCTAATTC 838
Dh 1258 aaggcaagaaagagccagctgagcttggcgaaagccaaagctaaagctgaagctatlc 1317
Oy 839 GAATCTGTGCTGAGCTGTGACACAAATATGAGATGAGATGAGATGAGATGAGATGAG 898
Dh 1318 gaaatctgtgctgagctctgacacaaataatgagagatgagcagctctcaagctgag 1377
Oy 899 CCGACGATGATGTACGCGCTTCTCCAACTGGCCAAAGACTTCACACTATCTACTGCT 958
Dh 1378 ccgagcagatgtgcagcggtcttccaactgccaagagactccaacactatctactgc 1437
Oy 959 CCTCAACCTGGCGATGTCACAGATGATGCTGAGGCTGAGGCTGATGATGAGAGCC 1018
Dh 1438 cctcaaacctggcgatgtaccacagcaatgtgctgagcacaatggtgtatataagagccc 1497
Oy 1019 TCACCAAAAGCCAGTGCAGGAGCTCGAGCTCACTCTCAGTGGGAGACAGAGATG 1078
Dh 1498 taccacaagcccaatgtccaagagactccagactcaactccagtgaggagcagcagagatg 1557
Oy 1079 TCCAGGATACAGATGCAACTGATGAGAACTTGTATGAGTCAAGATGAGTAACTGGA 1138
Dh 1558 tccaggtacagatgcaagctctgtagaggaactgtatcgatcaagatgagtagtgga 1617
Oy 1139 GCTGGCTTNGCCAGGAGTCTGGGAGCAAGAGCAAGATTTCCGATT 1188
Dh 1618 gctgggcttgccagggagcttgaggagcaagagcagatttccctgatt 1667

RESULT 5
US-10-219-793-45

Sequence 45, Application US/10219793
GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 86 Human Secreted Proteins

FILE REFERENCE: P2008P1

CURRENT APPLICATION NUMBER: US/10/219,793

PRIOR FILING DATE: 2002-08-16

PRIOR APPLICATION NUMBER: 09/209,462

PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45
LENGTH: 1337
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1335)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1336)
OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-219-793-45

```
Query Match          96.3%: Score 1144.4; DB 6; Length 1337;
Best Local Similarity 99.4%: Pred. No. 1,1e-276;
Matches 1167; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 16 CCGCTCCGCTCTCTGTTGTTCCGAGTCGTCGGCGGTGGGAAATGCTGGCGGC 75
Db 28 cgtctccgtctcgtctgtctcgtcggagtcgtcgtcgtcgtcgtcgtcgtcgtc 87
Oy 76 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 135
Db 88 gggggggc-ggggcaactgggccccttctgtgaggggtctctactgtctctcgtc 146
Oy 136 TCCG-GCCGCGCTCTCTGATGATGCGGAAACCGCTGATCTGTCGCGCAGCA 194
Db 147 tccgcgccgctcctcctcctcgtatgctcccgaaacacgtgtactgtctgtcggagca 206
Oy 195 GGAGCCCTGGGTGTGAGACGAATGGCGCGATCCACCGGATCTGAGACCTGTTGAA 254
Db 207 ggaagccctgggtgtgtgagcgaatggccgattccacccggtcctgtgagcctgtgaa 266
Oy 255 CATCTCATCCCTGTGTTAGACCGGATCGGATATGTGCGAGATCTCAAGAAATGTCAT 314
Db 267 catctcataccctgtgtatgagaccggatccgatatgtcagaatctcaagaatctgtcat 326
Oy 315 CAACGTGCTGAGACGATCGGCTGTGACTCTCGACAATGATTAAGTCTGCAATGATGAGAT 374
Db 327 caagctgtcctgtgagcagtcgtcgtctgtactctcgaatagtactctcgaatgtgagct 386
Oy 375 CCTTACTCGCCATCATGAGACCTTACAGCAGACTACGGTGTGAGAGACCTGTAGTA 434
Db 387 ccttactcgtcgtatcgtacgtacccttaccaggaagctacggtgtgtgagaccctgtgaa 446
Oy 435 TGCCGTGACCCAGTACGCTCAACAACATGATGATCAGAGTCGCGCAACTCTCTTTGA 494
Db 447 tgcgttcaaccgctcgtacgtccaaacaacatgtatcagaagctcgtcgaaactctctgtga 506
Oy 495 CAAAGTCTTCCGGGAACGGGAGTCCCTGATGCCAGCATTTGTGATGCCATCAACCAAC 554
Db 507 caaagtcttccgggaacgggagctccctgtatgtccagcatctgtgagtccatacaacgaagc 566
Oy 555 TGCTGACTGTGGGGTATCCGCTGCTNGTGTATGATGATCAAGGATATTCATGTGCCACC 614
Db 567 tgcgtactgtcgtgggtatccgtcgtcgttactgtatgatacgaagatacctatgtgcacc 626
Oy 615 CCGGGTGAAGAGTGTATGACAGATGACAGTGTGAGGCGAGCGGCGGAACGGGCGCCAGT 674
Db 627 ccgggtgaaagagctctatgtcagatgtgaggtcgaagcggcgagaaacgggcccacgt 686
Oy 675 TCTAGAGTCTGAGGGACCCGAGAGTGGCCATCAATGTGCGAGAGGAGAAACAGGC 734
Db 687 tctagagctctgagggagaccggagctcggtccatcaatgtgtagaaggaagaaacaagc 746
Oy 735 CCAGATCTCGGCTCGGAAGCAAGAAAGGCTGACAGATAAATACGGCGAGGAGAGGC 794
Db 747 ccagatctcggctcccgagacagaagaagctgtgaacagataaalcaggcgaggaagagc 806
Oy 795 CAGTGCATTTCTGGCGAAGGCGCAAGGCTAAAGCTGAAGCTATTCGATCTGCTGCAGC 854
Db 807 cagtgcatgtctcgtcggagagccaaagctaaagctgaagctatctgaatctcgtcgtcagc 866
Oy 855 TCTGACACAACTAATGAGATGACAGACGCTTCACTGACTGTGCGCGGAGCACTATGTACG 914
Db 867 tctgacacaacaataatgtgagctcagcagcttctactgtcgtcgtcgtcgtcgtcgtcgtc 926
Oy 915 CGCGTTTCCAAATGCGCCAAAGACTCCAAACTATCTACTGCGCTTCCAAACCTGTGGCA 974
|||||
```

```
Db 927 cgtcttccaaactgtgcgaagactccaactatccctactgtccctccaccctgtgcga 986
Oy 975 TGTMCACAGCATGTCGTCACGACGCAATGGGTATATGAGCGCTACCAAGCCCAT 1034
Db 987 tgtcacagcatgtgtgcacagcaltgtgtatgtagaccctcaacaagccccagct 1046
Oy 1035 GCCAGGACTCCAGACTCTCCAGTGGGAGCAGACAGATGTCACGATACAGATGC 1094
Db 1047 gccagagactccagactcctctccagtggtgagcagcaagatgtccaggtacagatgc 1106
Oy 1095 AAGTNTGATGAGGAACCTTGATCAGTCAAGATGACTTGTGAGAGCTGGCTTTGCCACG 1154
Db 1107 aagctctgtgtggaactgtatcagatgaatgaatgtgagctgtggtcgtgcag 1166
Oy 1155 GAGTCTGGGACAGAGACGATTTCTGTGATT 1188
Db 1167 gagctcgggacaagaagacagatcttccctgact 1200
```

RESULT 6

US-09-629-469A-11009
; Sequence 11009, Application US/09629469A

GENERAL INFORMATION:

APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629,469A
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11009
LENGTH: 1244
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (35)..(1102)
US-09-629-469A-11009

```
Query Match          95.2%: Score 1130.4; DB 5; Length 1244;
Best Local Similarity 99.4%: Pred. No. 3,9e-276;
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 30 TCGTTGGTTCCGAGAGTCTCTGCGGCTGTGGAAATGCTGCGCGCGCGCGGCGCA 89
Db 1 tctgtgttcgaggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 99
Oy 90 CTGGGGCCCTTTTGTGAGAGGGCTCTACTGTGCTTGTGGCGCGCTCCG-GCCGCGCCT 148
|||||
```

Db 60 ctggggcccttctgtaggggctctctactctgctctctgctccgctccgcgcgcgcct 119
Oy 149 CCTGTGATTTGGCCGAAACACCGTGTACTGTCTGTCGGCAGCAGGAGGCTGGGTG 208
Db 120 cctctgtagtgcgggaaacacgcgtgctactctgtctgcgcgagcagggcctgggtg 179
Oy 209 TGGAGCGAATGGGCGGATTCACCGGATCTTGAGCCTGGTTGAACATCTATCCCTG 268
Db 180 tggagcgaaatggccgattccaccgattccctggagcctgtgttgaaacatccatccctc 239
Oy 269 TGTTAGACCGGATTCGATATGTGACAGGTCTCAGGAATTTGTCTCAACGTCCTGAC 328
Db 240 tgltagaccggatccgatactgtagagctcaaggaaatgtlcatcaacgtgctcgaagc 299
Oy 329 AGTCGGCTGTGACTCTCGACATGTAACTCTGCAATTCAGATGAGTGGAGTCTTACCTG 388
Db 300 agtcggctgtgactctcgacaaatgtaactctcgaaatcgatggatctctactctgcga 359
Oy 389 TCATGACCCCTTACAAGCGACCTACGGTGTGAGGACCCCTGATATGCCCTCACCCAC 448
Db 360 tcatgacccttacaaggaaagctacggtgtgaggagaccctgatatgtccctcacccagc 419
Oy 449 TAGCTCAAAACCAATGAGATCAGAGCTCGGCAACTCTCTTTGACAAAGTCTTCGGG 508
Db 420 tagctcaaaacacacatgagatcagagctcgcaaacctctcttgacaagaatcttcgg 479
Oy 509 AACGGGAGTCCCTGAATGCGACATTTGGATGCCATCAACCAAGCTGCTGATCTGGG 568
Db 480 aacgggagctccctgaatgacgacatctgtagatgacatcaaacacgctgctgctgctg 539
Oy 569 GTATCCGCTGCTTNCCTTATGAGATCAAGATATCCATGTGCGACCCCGGTGAAGACT 628
Db 540 gttccgcgtcctcgttatgtagatcaagagatatacattgtgcaccccggtgaaagatc 599
Oy 629 CTATGCAATGAGGTGGAGGCGAGCGCGGAACGGCGCCACAGTTCTAGAGTCTGAGG 688
Db 600 ctatgcagatgtaggttggagcagagcgagcaagcgccacacatctctajagatcgaag 659
Oy 689 GGACCCGAGATCGGCGCATGATGCGAGGAGGGAAGGAACAGGCGCCGATCTGCGCT 748
Db 660 ggaacccggaggtcgagccatcaatgtgagcagaaggaaagaaacagcgccagatccctg 719
Oy 749 CGGAAGCAGAAAAAGCTGAACAGATTAATCAGCAGCAGAGAGGCCAGTGCAGTTCTG 808
Db 720 ccgaagcagaagaagctgtaacagataaatacagcagcagagagagcgagctgacgttc 779
Oy 809 CGAAGGCCAAGCTTAAGCTGAAGCTATTGAACTCTGCTGCTGACGCTTGACACAACAT 868
Db 780 cgaagcgcaagcttaagctgaagctatcgaatcttgctgacgtctgacacaaacata 839
Oy 869 ATGAGATGCACACGCTTACTGATGATGTCGGCAGAGAGATGTCAGCGGCTTCTCCAAC 928
Db 840 atgagagatgcaagcagcttactgactgtgtgcccgaagcagatagtctcagcgctctcca 899
Oy 929 TGGCCAAAGACTCAACACTATCTACTGCTCCCAACCCCTGGCATGTCTACACAGCATG 988
Db 900 tggccaagagactccaactatctctacgtgcccctccaacccctgscgtatgcaacagatg 959
Oy 989 TGGCTCAGGCGCTTGGGTGTATGTGAGCCCTTACCAAAAGCCCAAGTGCCHGAGACTCC 1048
Db 960 tggctcagcgctggtgtatatatgagccctcaacaaagcccaagctgcccgggagactcc 1019
Oy 1049 ACTCACTCTCCAGTGGAGCAGCAGAGATGTCAGGGTGCAGATGCAAGTNTTGTAGAG 1108
Db 1020 actcaactctccaggtggagcagagagatgtccagaggtacaaatgtaagccttgatgag 1079
Oy 1109 AACTGTATGACTCAAGATGAGTATAGTGGAGCTGGCTTGGCCAGGAGATCTGGGACAA 1168
Db 1080 aactgtatgagctcaagatgaggttagtggagcttgagccttgccagagagctctcgggag 1139
Oy 1169 GGAACGAGATTTTCTGATTT 1188
|||||

Db 1140 ggaagcagatttctctgalt 1159
RESULT 7
US-10-119-428-37
Sequence 37, Application US/10119428
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Wehrman, Tom
APPLICANT: Ren, Feiyang
APPLICANT: Ma, Yuning
APPLICANT: Zhou, Ping
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 789CIP2
CURRENT APPLICATION NUMBER: US/10/119,428
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/596,193
PRIOR FILING DATE: 2000-06-17
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PL-FL-genes Version 1.0
SEQ ID NO 37
LENGTH: 9098
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (75)..(9098)
US-10-119-428-37
Query Match 93.0%; Score 1104.8; DB 7; Length 9098;
Best Local Similarity 99.3%; Pred. No. 1.7e-269;
Matches 1128; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
Oy 1 GGCTTGTGGAGCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 12 ggccttggaagcagacgctcgctcgctcgctcgctcgctcgctcgctcgctcgctcgct 71
Oy 61 GAATGCTGCTCTACT 120
Db 72 gaaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 130
Oy 121 GGGTTGTGCTACT 179
Db 131 gggcttgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 190
Oy 180 GTTCGTCGCT 239
Db 191 gttcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 250
Oy 240 GGAGCTGCTTGAACATCTCATCTCTGTTAGACCGGATCCGATATGTGACAGTCT 299
Db 251 ggaagcctgttgtaaatctctatctctctctctctctctctctctctctctctctctct 310
Oy 300 CAAGGAATTTGATCATCAGCTGCTGAGCAGTGGCTGTGACTCTTGACATGTAATCT 359
Db 311 caaggaattgtcatcaacagttccttgagcagctgctgactctgacacaaatgtaacct 370
Oy 360 GCAAAATGATGAGTCTTACTTACCTGCGCATGATGAGCCCTTCAAGGCAAGCTACGCT 419
Db 371 gcaaatgtagagttcttactctgcatctgcatctgaccccttacaagcagctgagct 430
|||||

QY	420	GGAGGACCCCTGAGTATGGCCGTACCCAGCTAGCTAAACCAACATGACATCAGAGCTGG	479
Db	431	ggaggagccctgagtagtcgcgttaccagctagctctaaacacacatgagatcagagctcgg	490
QY	480	CAAACTCTCTNTGCAAAAGTTCTCCGGGAACGGGAGTCCCGAATGCCAGACTTGTCGA	539
Db	491	caaaactctctctcggacaagctcttcgcggaaaggaagctccctgaaatgccaagatctgga	550
QY	540	TGCCATTCACCAAGCTGCTGACTGCTGGGGTATCCGCTGCCCTMGTTATGAGATCAAGA	599
Db	551	tgcctaccacaagctctgactgctcgtcgtggtatccgcgcgcctccgctctatgatalcaaga	610
QY	600	TATCATGTGCGCACCCCGGTGAAGAGCTATGACATGTCAGATGCGAGTGGAGCGACGGCG	659
Db	611	tatcatactgcaaccccggtggaagaagctctatgagatgcagatggagagccagagcg	670
QY	660	GAACCGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGAGTGGCCATCAATGTGGCAGA	719
Db	671	gaaacgggcccacagctctcagagctcgtgagggagcccgagatcgcgcatcaatctgycaga	730
QY	720	AGGGGAAGAAACAGGCCAGATCTTGSGCTCCGAAACACAAAAGCTTGAACGATAAATCA	779
Db	731	aggggaagaaacagggcccagatccctgacctccgaagcagaagaagctcgaacagataaacca	790
QY	780	GGCAGCAGAGAGAGGCCAATGCGATTCTGTGCGAAGGCCAAGGCTAAAGCTGAAGCTATTCC	839
Db	791	ggcagcagagagagggccagctgcagctctcgtcgaaaggccaaggtctaaagctcgaagctatccg	850
QY	840	AATCCTGGCTCAGAGCTGTGACACACATTAATGAGAGATACAGAGCTTACAGACATGTCGC	899
Db	851	aatcctggctcagagctcgtgacacacacataagagaaatgcagaagactctacgtgacgtggc	910
QY	900	CGAGCAGTATGTACAGCGGCTTCTCCAACTGTGCGCAAGGACTTCAACACTATTCCTACTGCC	959
Db	911	cgaagcagatgctcagcgggtctctccaactcgtccaagagactccaacactatctctaagc	970
QY	960	CTCCAAACCTCGGATGTACACAGATGATGGTCTAGGCGCATGGGTGTATATGAGGCCCT	1019
Db	971	ctccaaacctcggcagctcgtgacacagcaatggctcgaagcctatggtatataagggcct	1030
QY	1020	CACCAAAAGCCCGACAGTGCACAGGACTTCACAGACTCATCTTCCAGTGGGACAGCAGAGAGT	1079
Db	1031	caccaaagcccccagatgcacagagatctccagactcctccagctggagcagagagatgct	1090
QY	1080	CCAGGGTACAGATGCAAGTNTTGATGAGAGAACTTATGAGTCAAGATGACTTAET	1135
Db	1091	ccagggtacagatgcaagctctctgataaggaactctgatactatgatacaagatgacttgct	1146

```

RESULT      8
US-09-705-256A-3468
; Sequence 3468, Application US/09705256A
; GENERAL INFORMATION:
; APPLICANT: Ma, Xiao-Jun; Dotson, Stanton B.; Monsanto Company
; TITLE OF INVENTION: Tumor Associated Molecules (TAMs); Targets for diagnosis, treatment
; TITLE OF INVENTION: and prevention of cancer
; FILE REFERENCE: 3214
; CURRENT APPLICATION NUMBER: US/09/705, 256A
; CURRENT FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: US 60/164, 285
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 3468
; LENGTH: 1234
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-705-256A-3468

Query Match      91.4%; Score 1086.4; DB: 5; Length 1234;
Best Local Similarity 99.2%; Pred. NO. 5.4e-265;
Matches 1131; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

QY	52	CGGGGTGGGAAATGCTGGCGCGCGCGGGGCGACATGGAGCCCTTTTGCTGAGGG	111
Db	1	cggcggtgggaagaagctgagcgcgcgcgcg-999agcattgggccccttctgtcgaagg	59
QY	112	CTCTCTACTGGCTCTCGGCGGCGCTCG-CCGCGCGCTCTCTGGATTGGCCCGAAAC	170
Db	60	ctctctactggtcttcgagcgcgctcgcgcgcgctctcttgattgcccgaagac	119
QY	171	CGTGGTACTGTTTCGTGGCGGAGAGAGGCTT-GGGTGTGGAGGAAATGGGCCATTTCC	229
Db	120	cgltgtaactgtctgtcgccagcgagagagcgcttgggtgtgtgagcgaaatggccgatcttc	179
QY	230	ACCGGATCTTGAGACCTGGTTTGAACATCTCATCCCTGTGTTAAGCCGATCCGATATG	289
Db	180	accggaatcccgagccgtgtgtgaactctcactctccctgtgttagaccggaatccgaatag	239
QY	290	TGCAGAGCTCAAGGAATTGTGATCATAGCTGCTTAGAGTTCGGCTGTGACTCTCGACA	349
Db	240	tgcagagctcaagaagaaattgtcaatacagctgcgtgagcagctggtctgtgaactctgcaca	299
QY	350	ATGTAACTCTCAATATGATGAGATGCTTTACTTGGCGGATCATGTGACCCCTTAAGGCAA	409
Db	300	atgtaaactctgcaaatctgaatgaggtcccttactctgtgcatctgtgaacctcaagaagcaa	359
QY	410	GCTACGGGTGTGAGACCTTGAGTATGCGCTGATCCAGCTAGCTCAAACAAACATGAGAT	469
Db	360	gctaagggtgtgagagaccctgagatagtcgltcaaccagctagctcaaaacaacatgagat	419
QY	470	CAGAGCTGGGCAACTCTGTTNTGGACAAGATCTTGGGGAGAGGGAGTCCCGAATGGCA	529
Db	420	cagagcttgggcaaacctctctctgcgaacaagctctccgggaacgggaagctccgtaatgcaca	479
QY	530	GCATTGTGGAGTCCCATACCAAGAGCTGACTGCTGGGGGTATCCGCTGCTTCATTATG	589
Db	480	gcattgtgtagatgccatacaacaagctgtgactgctgcgtgggtatccgctgcctcgltag	539
QY	590	AGATCAAGGATATTCATGTGTCACCCCGGCTGAAGAAGTCTATGACAGTCAAGTGGAGG	649
Db	540	agataaagagatatacctgtgcaccccggtgtgaagaagcttatgcaagatgcaggttgaagg	599
QY	650	CAGAGCGCGGAACGGGCGACAGTTCTAGAGCTGA-GGGAGCCGAGAGTGGGCCCATC	708
Db	600	cagagcgcggaagaacgggccaagcttcttaagctctgaagggggaacccgaagatcgcgcaatc	659
QY	709	AATGTGGCAGAGAGGAAGAAACAGGGCCCAANTCTTGCCCTCCGAAGCAGAAAAGCTTAA	768
Db	660	aatgtggcagaagaaggaaagaacaggccccgaatctctgctccggaagcagaaagaagcttgaa	719
QY	769	CAGATTAATCAGGACGACGAGAGAGGCCAGTGTGAGTTCTTGCGCAAGGCCAAGGCTAAAGCT	828
Db	720	cagataaatcagcgacgacgaaggaaggccaggtgcagttctgtgcgaaggccaaagctaaagct	779
QY	829	GAAGTATTCGAATCTTGCGCTGCAGCTCTTGACACACATATATGAGATGTCAGAGCTTCA	888
Db	780	gaagtatcttgaatctccgtgcagctctgtgacaaacataatgtgagatgtgcagcagcttcaa	839
QY	889	CTGACTGTGGCGGAGCAGATATGTGACGGGCTTCTCCAAACTGGCCAAAGACTCCCAACT	948
Db	840	ctgactgtggcgcgagcagtatgtcagcggtcttccaaactgtgcccgaagactccaacact	899
QY	949	ATCTACTGCCCCCAACCCCTGGCGATGTACACGAGATGGTGGCTCAGGCCATGGGTGTA	1008
Db	900	atctactgcccctccaacccctgtgcgagtgcatacagaatggtgtcagtgccatgggtgtgta	959
QY	1009	TATGGAGCCCTCACCAAGGCCCAAGTGCACAGGAGCTTCAGACTCACTCTCCAGTGGAGC	1068
Db	960	tatggagccctcaccaagggcccaagcttcagaggatctcagactcaactctccagtgaggac	1019
QY	1069	AGCAGAGATGTCACGGGTACAGATGCAAGTNTGATGAGAGAACTTGATGAGTCAAGTAAAGT	1128
Db	1020	agcagaagatgtccagggttacaagaatgcagatctgtgaatgaacttgaatcgagatcaagatc	1079
QY	1129	AGTTAGTGGACTTGGGCTTNNCCAGGGAGTCTTGGGACAAGGAACACAGATTTTCTTGATT	1188

OY	378	TTAACTCGCATATGAGACCTTACAAAGCAAGCTACGGTGTGGAGGACCCGATATAGC	437
Db	167	ttaccgtcgacatacaggaacccctacaaggcaagctacaggtgtgagaggaaccctgaatagc	226
OY	438	CGTCCACCACGCTAGCTCAACAAACATAGAGATCAGAGCTCGGCAAACTCTCTNTGGACAA	497
Db	227	cgtaaccacagctagctcaaaacaacccaatgatacagagcttcggaacaatctctctgaaaca	286
OY	498	AGCTCTCCGGGAACGGGAGTCCCTGAATGCCAGACATTTGTGATGCCATCAACAAAGTGC	557
Db	287	agctctccgggaacggagagctccctgaatgcagcatctgtagatgcacataacaaagctgc	346
OY	558	TGAATGCTGGGGTATCCGGCGCTTCGTTATAGATGACAGGATATCAATGTGCCACCCCG	617
Db	347	tgaaatgctggggtatccggcgcttcggttatagatgacaggaatatcaatgtgccaccg	406
OY	618	GGTGAAGAGACTATATGAGATTCAGCTGGAGAGCAGACGGCGGAACGGGACACAGTCT	677
Db	407	ggtagaagagatctatcagatgcaagcttgagagcgagagcgcggaacgggccaagcttct	466
OY	678	AGACTGTAGAGGGACCCGAGAGTGGCCATCAATGTGCAGAAAGGAAAGAAACAGGCCCA	737
Db	467	agagctcagagggagcccgagagctcgagcatalcaatgtgcagaaggaagaaacagcccca	526
OY	738	GATCTCTGGCTCTCGAAGAGCAAAAGGCTGAACAGATAAATCAGGACAGCAGAGAGCCAG	797
Db	527	gattctctgctctcgaaagcagaataaaggctgaacaagataaaatcagcagaagagagccag	586
OY	798	TGCAGTTCTTGCGCAAGGCCMAAGGCTTAAAGCTTAAGCTATTCGAATCTGGCTGACGCT	857
Db	587	tgcagttcttgcgcaagggccaaagcttaagctgaagctatctgaaatccctgcgctcgaa	646
OY	858	GACACAACATATATGAGATGCAGACAGCTTCACTGACTGTGGCCGAGCAGATATGTACAGGC	917
Db	647	gacacaacataatgagatgcagagcttcaactgactgtggccgagcagtatgtcagcgc	706
OY	918	GTTCTTCCAAACTGGCCCAAGAGACTCTCAACACTATCTACTGCTCCCTTCAACCTCGGAGATGT	977
Db	707	gttcttccaaactggcccaagagcttcaacactatctactgctccctcaaacctctgtagatgt	766
OY	978	CACCAGATGGTGGCTCAGAGCCATATGGGTATATGAGAGCCCTCACCAAGCCCGAGTGC	1037
Db	767	caccagatggctggctcagagccatattgggtatatgagagccctcacaaagcccgagctgc	826
OY	1038	AGGAGACTCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCACGGGTACAGATGCAAG	1097
Db	827	a-ggaactcagactcaactctccagtgaggagcagcagagatgtcca-gtlaacagatgcaag	884
OY	1098	TNTTGTAGAGAGAACTTGATGAGTCAAGATGAGTTAGTGGAGCTGG---CTTNGCCAGG	1154
Db	885	tctgtagaagaaacttgatctgaatcaagaatgagcttagtgagacgtctgagcttgagccag	944
OY	1155	GAGTCTGGGAGACAAGGACAGATTTTCCGTGATT	1188
Db	945	gagctcgggagacaagagcagatcttcccgatt	978
RESULT 11			
PCT-US01-43704-1161/C			
: Sequence 1161, Application PC/TUS0143704			
: GENERAL INFORMATION:			
: APPLICANT: Corixa Corporation			
: APPLICANT: Stolk, John A.			
: APPLICANT: Xu, Jiangchun			
: APPLICANT: Chenualt, Ruth A.			
: APPLICANT: Meagher, Madelein Joy			
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND			
: FILE REFERENCE: 210121.561PC			
: CURRENT APPLICATION NUMBER: PCT/US01/43704			
: CURRENT FILING DATE: 2001-11-19			
: NUMBER OF SEQ. ID NOS: 2606			
SOFTWARE: Corixa Invention Disclosure Database			

```

: SEQ ID NO 1161
: LENGTH: 622
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 615
: OTHER INFORMATION: n = A,T,C or G
: CDT:OS01-43704-1161

```

Query Match	52.1%;	Score 618.6;	DB 1;	Length 622;
Best Local Similarity	99.4%;	Pred. No. 1.1e-146;		
Matches 618; Conservative	1;	Mismatches 3;	Indels 0;	Gaps 0;

OY	310	GTATCAACGTGACCCGAGAGAGTGGCGCTGAGACTCTCGACAAATGTAACCTGTGAAATTCGAT	369
Db	622	GTTCATCAACGKACCCCTGAGAGAGTGGCGCTGAGACTCTCGACAAATGTAACCTGTGAAATTCGAT	563
OY	370	GGAGTCTCTTTACCTCGCATCATGAGACCCCTTACAAAGGCAAGCTACGGTGTGAGAGACCT	429
Db	562	GGAGTCTCTTTACCTCGCATCATGAGACCCCTTACAAAGGCAAGCTACGGTGTGAGAGACCT	503
OY	430	GAGTATGCCCGTCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGGCAAACTCTCT	489
Db	502	GAGTATGCCCGTCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGGCAAACTCTCT	443
OY	490	NTGCAACAAAGTCTCCGGGAACGGGAGTCCCTGAAATGCCAGATGTGGATGCCCATCAAC	549
Db	442	CTGGCAAAAGTCTTCGGGAAACGGGAGTCCCTGAAATGCCAGATGTGGATGCCCATCAAC	383
OY	550	CAAGCTGCTGACTGCTGGGGGTATCCGCTGCTTCGCTTATGAGATCAAGGATATTCATGTG	609
Db	382	CAAGCTGCTGACTGCTGGGGGTATCCGCTGCTTCGCTTATGAGATCAAGGATATTCATGTG	323
OY	610	CCACCCCGGGTGAAGAAGTCTATGCAGATGCAGGTGGAGGACAGAGCGCGGAAACGGGCC	669
Db	322	CCACCCCGGGTGAAGAAGTCTATGCAGATGCAGGTGGAGGACAGAGCGCGGAAACGGGCC	263
OY	670	ACAGTTTAAAGATCTGAGGGGGACCCGAGAGTGGCGCATCAATGTGGCAAGAGGAAAGAA	729
Db	262	ACAGTTTAAAGATCTGAGGGGGACCCGAGAGTGGCGCATCAATGTGGCAAGAGGAAAGAA	203
OY	730	CAGGCCACAGATCTTGCGCTCCCGAACAGAGAAAGGCTGAACAGATTAATCAGGCAGCAGGA	789
Db	202	CAGGCCACAGATCTTGCGCTCCCGAACAGAGAAAGGCTGAACAGATTAATCAGGCAGCAGGA	143
OY	790	GAGGCCAGTGCAGTTCTTGCGCGAAGGCCAAAGGCTAAAGCTGAAGCTATTTCCTGGCT	849
Db	142	GAGGCCAGTGCAGTTCTTGCGCGAAGGCCAAAGGCTAAAGCTATTTCCTGGCT	83
OY	850	GCAGCTCTGACACACATTAATGAGATGAGAGCTTCACCTGACACTGGCCGACAGCAGTAT	909
Db	82	GCAGCTCTGACACACATTAATGAGATGAGAGCTTCACCTGACACTGGCCGACAGCAGTAT	23
OY	910	GTCAGCGCGCTTCTCCAAACTGG 931	
Db	22	GTCAGCGCGCTTCTCCAAACTGG 1	
RESULT 12			
US-10-214-403-2995/c			
: Sequence 2995, Application US/10214403			
: GENERAL INFORMATION:			
: APPLICANT: Jiang, Yugu			
: APPLICANT: Chenault, Ruth A.			
: APPLICANT: Xu, Jiangchun			
: APPLICANT: Indrias, Carol, Yoseph			
: APPLICANT: Lodes, Michael J.			
: APPLICANT: Secrist, Heather			
: APPLICANT: Carter, Darrick			
: APPLICANT: Fanger, Gary R.			
: APPLICANT: Smith, Carole L.			


```

: CURRENT FILING DATE: 2002-03-13
: NUMBER OF SEQ ID NOS: 1562
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1428
: LENGTH: 691
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 439, 508, 523, 526, 539, 561, 575, 582, 584, 602, 614, 623,
: LOCATION: 626, 645, 648, 664
: OTHER INFORMATION: n = A,T,C or G
US-10-097-105-1428
```

```

Query Match          43.5%: Score 516.6; DB 7; Length 691;
Best Local Similarity 91.4%: Pred. No. 7,4e-121;
Matches 616; Conservative 0; Mismatches 46; Indels 12; Gaps 7;
```

```

QY 200 CTTGGTGTGAGGAGGAGTGGCCGATTCACCCGATCCTGAGCCCTGTTGAACATCC 259
D 1 cctgggtgtgtgaggaatgggcccgaatccaccggaatccttgagccttggtgaacatcc 60
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAGAGAAATTGTCAACAAG 319
D 61 tcatccctgtgttagaccggatccgatatgtgcagagttcagaagaaatgtcatcaacg 120
QY 320 TGCCTGACAGCTCGGCTGTGACTCTGACAAATGTACTCTGCAATGATGGAGTCTTT 379
D 121 tgcctgagcagatcgctgtgaactctcgacaatgttaactctgacaatcgatggatcctt 180
QY 380 ACCTGCGCATATGAGACCCCTTACAAGGCAAGCTACGCTGTGGAGAGCCCTGAGTATGCCG 439
D 181 acctgcatcatatgagacccttacaagcaagctacgctgtggagagccctgatatgccc 240
QY 440 TCACCCAGCTAGCTCAACCAACCATGAGATCAGAGCTGGCAAACTCTCTNTGACAAAG 499
D 241 tcaccagctagctcaacaacccaatgagatcagagctcgacaactctcttcgagcaaaag 300
QY 500 TCTTCGCGGAGACGGAGATCCCTGATGCGCAGATGTCATGATGTCATCAACCAAGCTGCTG 559
D 301 tcttcgsgaagcgagagtcctcctgaaatcgacagatgtgagatgcatcaacaagctgtg 360
QY 560 ACTGCTGGGGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
D 361 actgctgggtatccgctcgcctcgtlatagagatcaaggaatcatalgtgccaccg 420
QY 619 GTGAAG-AGTCTATGAGATGAGTGGAGGAGAGC-GGCGGAAAGGGCCACAGTTC 676
D 421 gtgaagaagctcatalgcanatgagatgagatgagatgagatgagatgagatgagatg 480
QY 677 TAGAGTCTGAGGAGCAGCCGAGAGTGGCCATCATGTCGACAGAG- - - - -GAGAGACA 731
D 481 tagagctcgaaggagagccgaaagtcgncatcatalgtgsgaanaaggagaagaacang 540
QY 732 GGGCCAGATCTCGGCT- - -CCGAGACAGAAAAGCTGAACAGA-TAATCAGAGCAGAG 788
D 541 gggccagatctcggctcctcctcnaaacaagaagatgaacanaataaatacagaccg 600
QY 789 AG-AGCCAGTACTTGTGCGAGAGCCAGAGGCTAAAGCTAATGCTATTCGATCTGG 847
D 601 aagagccatgtgctgttcttgcaanaagcaagatcaaaaactgaanctnttgaaatcc 660
QY 848 CTGCAAGCTCTGACA 861
D 661 gtgnaactttacca 674
```

```

RESULT 15
US-09-629-469A-1950
: Sequence 1950, Application US/09629469A
: GENERAL INFORMATION:
: APPLICANT: OTA, TOSHIO
```

```

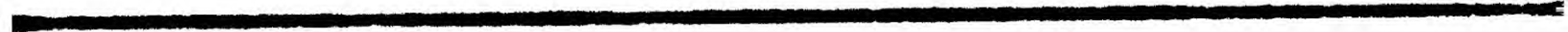
: APPLICANT: ISOGAI, TAKAO
: APPLICANT: NISHIKAWA, TETSUO
: APPLICANT: HAYASHI, KOJI
: APPLICANT: SATTO, KAORU
: APPLICANT: YAMAMOTO, JUNICHI
: APPLICANT: ISHII, SHIZUKO
: APPLICANT: SUGIYAMA, TOMOYASU
: APPLICANT: WAKAMATSU, AI
: APPLICANT: NAGAI, KEIICHI
: APPLICANT: OTSUKI, TETSUJI
: TITLE OR INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
: FILE REFERENCE: 084335/0123
: CURRENT APPLICATION NUMBER: US/09/629, 469A
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: JP 1999-248036
: PRIOR FILING DATE: 1999-07-29
: PRIOR APPLICATION NUMBER: JP 1999-300253
: PRIOR FILING DATE: 1999-08-27
: PRIOR APPLICATION NUMBER: JP 2000-118776
: PRIOR FILING DATE: 2000-01-11
: PRIOR APPLICATION NUMBER: JP 2000-183767
: PRIOR FILING DATE: 2000-05-02
: PRIOR APPLICATION NUMBER: JP 2000-241899
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/159,590
: PRIOR FILING DATE: 1999-10-18
: PRIOR APPLICATION NUMBER: 60/183,322
: PRIOR FILING DATE: 2000-02-17
: NUMBER OF SFO ID NOS: 19025
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1950
: LENGTH: 567
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (258)..(258)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (333)..(333)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (379)..(379)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (425)..(425)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (438)..(438)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (441)..(441)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (445)..(445)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (508)..(508)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (533)..(533)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (561)..(561)
```

```
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (563)..(563)
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (567)..(567)
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-09-629-469A-1950
```

```
Query Match          43.5%; Score 516.4; DB 5; Length 567;
Best Local Similarity 97.0%; Pred. No. 8e-121;
Matches 550; Conservative 0; Mismatches 14; Indels 3; Gaps 3;
```

```
OY 30 TCGTTGGTTCGAGAGTCCGTGCGCGGTGGGAAATGCTGGCGCGCGCGCGGGA 89
Db 1 tcgltgltccgaggtcgcgtggtggaatgctgagcgcgcgcgca 59
OY 90 CTGGGCGCTTTTGTGAGGGGCTCTACTAGCTTGTGGCGCGCTCCG -GGCGGCGCT 148
Db 60 ctggggcccttctgcgagggctctctactagcttctgcgcgcgcgcgct 119
OY 149 CCTCTGATTTGCCCGAACAACCCGTGTACTGTCTGTCGCGCAGCAGAGCCTGGGTG 208
Db 120 cctctgattgcccgaataaacggtgtactgtctgtccgcaagagcctggtg 179
OY 209 TGGAGCGAATGGCCGATTCACCGGATCTGAGCCTGGTTGAACATCTCATCCCTG 268
Db 180 tggagcgaaatggccgattccaccggtcctgagcctggtltgaaatcctcctc 239
OY 269 TGTTAGCCGGATCCGATATGTCAGAGTCTCAGGAAATGTATCAAGTGCCTGAGC 328
Db 240 tgttagaccggatccgactcgtgagagttccaaagaaatgtcatcaacgtgctgagc 299
OY 329 AGTCGGCTGTGACTCTGCACATGTACTCTGCAATCGATGAGTCCCTTACTGGCA 388
Db 300 agtcgctgtgactctcgacaatgtactctcgtacatgagtgagtccttacttgca 359
OY 389 TCATGGACCTTACAAAGCAGCTACGGTGTGAGAGCCTGAGTATGCCGTACCCAGC 448
Db 360 tcattgaccttacaacgcnagctacgtgttgagagacctgagtalgcgtcacccagc 419
OY 449 TAGCTCAACAACATGAGATTCAGAGCTCGGCAAACTCTCTNNGACAAATCTTCGGG 508
Db 420 tagctcaacaacaacatganaftagancctcgcaactctctctcgacaaatcttcg 479
OY 509 AAGGGAGTCCCTGAAATGCGAGCATTGTGGATGCCAT -CAACCAAGCTGCTGACTGG 567
Db 480 aacggagagtcctcgaaatgccaagcatgtgngatgccatccaacaagctgctgagt 539
OY 568 GGTATCCGCTGCTTCGTTATGAGATC 594
Db 540 ggtatccgctgcttcctcgltatnanatc 566
```

Search completed: September 22, 2002, 16:20:21
Job time: 5453 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:21:14 : Search time 1684.32 seconds

(without alignments)
9519.796 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCCTCTGGAGACNACCGCT.....GGAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues 27473414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthbm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1067.2	89.8	1218	11 AF161458	AF161458 Homo sapi
2	916.6	77.2	935	9 AL541792	AL541792 AL541792
3	899.4	75.7	1110	10 BM476304	BM476304 AGENCOURT
4	884.2	74.4	962	9 AL516503	AL516503 AL516503
5	871.8	73.4	994	10 BM468527	BM468527 AGENCOURT
6	860	72.4	998	10 BM468407	BM468407 AGENCOURT
7	842.2	70.9	1536	11 AK002793	AK002793 Mus muscu
8	835.4	70.3	1011	10 BM461614	BM461614 AGENCOURT
9	821.4	69.1	849	9 AL525168	AL525168 AGENCOURT
10	821	69.1	1061	10 BM453159	BM453159 AGENCOURT
11	802.6	67.6	899	9 AL525167	AL525167 AGENCOURT
12	792.6	66.7	1095	10 BM423347	BM423347 AGENCOURT
13	791.8	66.6	918	10 BM459766	BM459766 AGENCOURT
14	791.4	66.6	943	10 BM479080	BM479080 AGENCOURT
15	791.2	66.6	920	10 BM462105	BM462105 AGENCOURT
16	791	66.6	925	10 BM43426	BM43426 AGENCOURT
17	790.8	66.6	820	10 BM685336	BM685336 AGENCOURT

C	18	770.4	64.8	787	9	AL569086
	19	766.2	64.5	796	10	BI092184
	20	764.4	64.3	1058	10	BM423453
	21	762.8	64.2	1053	10	BM456748
	22	757.4	63.8	884	10	BM456611
	23	757.4	63.8	887	10	BM045426
	24	755.6	63.3	923	10	BI870704
	25	752.4	63.3	853	10	BE314702
	26	750.4	63.2	988	10	BM048505
	27	749	63.0	800	10	BM45196
	28	743.4	62.6	836	10	BM823189
	29	740.8	62.4	924	10	BI767424
	30	738	62.1	1017	10	BM753777
	31	736.2	62.0	778	10	BM681190
	32	724.6	61.0	809	10	BI086063
C	33	715.8	60.3	770	10	BM828070
	34	715.4	60.2	903	10	BM532231
	35	711.6	59.9	813	10	BM768715
	36	709.4	59.7	856	10	BM965844
	37	706.2	59.4	916	10	BI260046
	38	703.8	59.2	752	10	BI552983
	39	702	59.1	728	10	BI522493
	40	702	59.1	739	10	BM574270
	41	701.2	59.0	865	10	BM051182
	42	700	58.9	986	10	BI694744
	43	698.4	58.8	788	10	BM767007
	44	698.2	58.8	785	10	BM009759
	45	697.4	58.7	1190	10	BF026777

ALIGNMENTS

RESULT	1
AF161458	AF161458
LOCUS	Homo sapiens HSPC108 mRNA, complete cds.
DEFINITION	AF161458
ACCESSION	AF161458.1
VERSION	GI:6841439
KEYWORDS	HTC.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1218) Zhang, Q.H., Fan, H.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.W., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J. and Chen, Z.
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)
MEDLINE	20499367
REFERENCE	2 (bases 1 to 1218) Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.
AUTHORS	Human full length cDNA cloned from cd34+ stem cells
TITLE	Unpublished
JOURNAL	3 (bases 1 to 1218)
AUTHORS	Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China.
FEATURES	location/Qualifiers
source	1..1218 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CBCAPB12" /cell_type="cd34+ stem cells" /tissue_type="blood"

```

CDS
18. 1046
/codon_start=1
/product="HSPc108"
/protein_id="AAF29073.1"
/db_xref="GI:6841440"
/translation="MRGSLASGRACRASSGLPNTVLFVPOEAMVVERMREHRI
LEPRLNLIPLVDRIKRVOSLKEVINVPDASATLIDNVLIDGVLXLRIMDPKAS
YVEDPEYATVQLQRTTMRSELKLSLDKVFREHESLNASIVDAINQADCWGRCLR
YEKIDIVPRVKESSKOMOVEAERKRRTYLESGETRESAINVEGRKQADILASEA
KAEQINOAGASASVLAKAKAKAEAIRILAAALOHNGDAAASLTVAEOVYASKLK
KDSNTILLPSNPGDVTSMVAQMGVYGLATKAPVPTIPDLSLSSSSRDVQDTASLDE
ELDRVKMS"
BASE COUNT      313 a      310 c      337 g      258 t
ORIGIN
Query Match      89.8%; Score 1067.2; DB 11; Length 1218;
Best Local Similarity 99.1%; Pred. No. 1.2e-250;
Matches 1092; Conservative 0; Mismatches 7; Indels 3; Gaps 2;
QY 88 CACTGGGGCCCTTTTGC-TGAGGGGCTCTACTGCTTTCGGCCGCGCCGCGC 146
Db 1 CACTGGGGCCCTTTTGCATGAGGGGCTCTACTGCTTTCGGCCGCGCTT--GCCGCGC 58
QY 147 CTCCTCTGATTCGCCCAACACCGTGGTACTGTCGCGCGCAGCAGAGGCGCTG 206
Db 59 CTCCTCTGATTCGCCCAACACCGTGGTACTGTCGCGCGCAGCAGAGGCGCTG 118
QY 207 GGTGAGGCAATGGGCGCATTCACCGGATCCTGGAGCGCTGGTTGAACATCTCATCC 266
Db 119 GGTGAGGCAATGGGCGCATTCACCGGATCCTGGAGCGCTGGTTGAACATCTCATCC 178
QY 267 TGTGTAGACCGGATTCGATATGTGACAGTCTCAAGAAATTGTATCAACTGCTCTA 326
Db 179 TGTGTAGACCGGATTCGATATGTGACAGTCTCAAGAAATTGTATCAACTGCTCTA 238
QY 327 GGAGTCGGTGTGACTCTGCACATGTACTGCAAAATCGATGGATGCTTACTGCG 386
Db 239 GGAGTCGGTGTGACTCTGCACATGTACTGCAAAATCGATGGATGCTTACTGCG 298
QY 387 CATCATGACCCCTTCAAGAGCAAGCTACGCTGTGAGAGCCCTGAGTATCCGTCACCA 446
Db 299 CATCATGACCCCTTCAAGAGCAAGCTACGCTGTGAGAGCCCTGAGTATCCGTCACCA 358
QY 447 GCTAGCTCAACACCATGATGATGAGTCCGCAACTCTCTNTGGACAAGTCTTCGG 506
Db 359 GCTAGCTCAACACCATGATGATGAGTCCGCAACTCTCTCTGACAAAGTCTTCGG 418
QY 507 GGAACGGGAGCTCCCTGAATGCCAGCATGTGGATGCCATCAACCAAGCTGCTGCTG 566
Db 419 GGAACGGGAGCTCCCTGAATGCCAGCATGTGGATGCCATCAACCAAGCTGCTGCTG 478
QY 567 GGGTATCCCTGCTCCTGCTTATGATCAAGATATCATGTCGCCACCCCGGCTGAAGA 626
Db 479 GGGTATCCCTGCTCCTGCTTATGATCAAGATATCATGTCGCCACCCCGGCTGAAGA 538
QY 627 GTCTATGCAAGTCAAGTGGAGGCGGCGGAAACGGGCCACAGTCTTAGAGTCTGA 666
Db 539 GTCTATGCAAGTCAAGTGGAGGCGGCGGAAACGGGCCACAGTCTTAGAGTCTGA 598
QY 687 GGGGACCCCAAGATCCGCCATCAATGTGCGAGAAGGAAGAAACAGGCCAGATCCGCGC 746
Db 599 GGGGACCCCAAGATCCGCCATCAATGTGCGAGAAGGAAGAAACAGGCCAGATCCGCGC 658
QY 747 CTCGGAAGCAGAAAAGCTGTAACAGATTAATCAAGCAGCAGAGAGGCGCAGTCTT 806
Db 659 CTCGGAAGCAGAAAAGCTGTAACAGATTAATCAAGCAGCAGAGAGGCGCAGTCTT 718
QY 807 GCGGAAGGCGCAAGGCTAAAGCTATTTCAATCTGGCTGCAAGCTTGAACACACA 866
Db 719 GCGGAAGGCGCAAGGCTAAAGCTATTTCAATCTGGCTGCAAGCTTGAACACACA 778
QY 867 TAATGAGATGACGACGCTTCACTGACTGTGGCGGAGCAATATGTACAGCCGTTCTCAA 926

```

```

Db 779 TAATGAGATGACGACGCTTCACTGACTGTGGCGGAGAGATGATCACGGGTTCTCCA 838
QY 927 ACTGGCCAAAGACATTCACACTATCTTACTGCGCTCCAAACCTGGCGATGTCCACAGAT 986
Db 839 ACTGGCCAAAGACATTCACACTATCTTACTGCGCTCCAAACCTGGCGATGTCCACAGAT 898
QY 987 GGTGGCTCAGGCGATGGGTATATGAGAGCCCTCCCAAGGCCAGAGCCAGGACTCC 1046
Db 899 GTGGCTCAGGCGATGGGTATATGAGAGCCCTCCCAAGGCCAGAGCCAGGACTCC 958
QY 1047 AGACTCACTCTCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGTNTTGATGA 1106
Db 959 AGACTCACTCTCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGTNTTGATGA 1018
QY 1107 GGACTTGATGAGTCACAGATGATGATGAGAGCTGGCTTGGCCAGGAGCTGTGGGAC 1166
Db 1019 GGACTTGATGAGTCACAGATGATGATGAGAGCTGGCTTGGCCAGGAGCTGTGGGAC 1078
QY 1167 AAGGAAGCAGATTTTCTGATTT 1188
Db 1079 AAGGAAGCAGATTTCTGATTT 1100

RESULT 2
AL541792
LOCUS AL541792 935 bp mRNA linear EST 16-FEB-2001
DEFINITION AL541792 LTI_FL002_PL1 Homo sapiens CDNA clone CS0DE007YE18 5 prime
, mRNA sequence.
ACCESSION AL541792
VERSION AL541792.1 GI:12873201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..935
location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="CS0DE007YE18"
  /clone_1ib="LTI_FL002_PL1"
  /lab_host="DH10B"
  /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a Notti-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
technologies. Contact : Feng Liang Life Technologies, a
division of invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA fax : (1) 301 610 8371 Email :
filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      239 a      246 c      265 g      181 t      4 others
ORIGIN
Query Match      77.2%; Score 916.6; DB 9; Length 935;
Best Local Similarity 99.0%; Pred. No. 7.9e-214;
Matches 927; Conservative 3; Mismatches 5; Indels 1; Gaps 1;
QY 189 GCAGCAGAGAGCGCTGGTGTGAGAGCAATGGGCCGATTTCCACCGGATCTGAGGCTGG 248
Db 1 GCAGCAGAGAGCGCTGGTGTGAGAGCAATGGGCCGATTTCCACCGGATCTGAGGCTGG 60

```



```

Oy 575 GCTGCTTCGTTATGATGATCAAGATATCCATGTCGCCACCCCGGGTGAAAGATCTATCC 634
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 566 GCTGCTCCGTTATGATGATCAAGATATCCATGTCGCCACCCCGGGTGAAAGATCTATCC 625
Oy 635 AGATCAGGTGAGAGCAGACGCCGGGAAAGCGGCCACAGTCTTAGAGTGTGAGGGGACCC 694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 636 AGATCAGGTGAGAGCAGACGCCGGGAAAGCGGCCACAGTCTTAGAGTGTGAGGGGACCC 685
Oy 695 GAGATCGGCGCTCAATGATGGCAGAGGAAGGAAGAGCCAGATCTGGCCTCCGAAG 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 666 GAGATCGGCGCTCAATGATGGCAGAGGAAGGAAGAGCCAGATCTGGCCTCCGAAG 745
Oy 755 CAGAAAAGGCTGACAGATAAATCAGGAGCAGAGAGAGCCAGTGTGAGTGGCGAAG 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 746 CAGAAAAGGCTGACAGATAAATCAGGAGCAGAGAGAGCCAGTGTGAGTGGCGAAG 805
Oy 815 CCAAGGCTAAAGCTAAGCTATTCGAATCTGGCTGAGAGCTGTGACACAACTAATGAG 874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 806 CCAAGGCTAAAGCTAAGCTATTCGAATCTGGCTGAGAGCTGTGACACAACTAATGAG 865
Oy 875 ATGACAGAGCTTCACTGACTGTGGCGAGCAGTATGTGACGCG--TTCTCCAAAGTGGCC 933
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 ATGCGCGAGCTTCCCTGACTGTGGCGAGCAGTATGTGACGCGCTTCTCCAAAGTGGCC 925
Oy 934 AAGACTCCCAACACTATCTACTGCTCCCAACCTGGCG--ATGTCACAGCAGATGGTGC 992
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 AAGACTCCCAACACTATCTACTGCTCCCAACCTGGCGAGGCGCCAGCAGATGGGAGC 985
Oy 993 TCA--GGCATGTGGTGTATATGAGGCCCTCACC---AAAGCCCCAGTGGCAGGACTCCA 1047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 TCAAGGCCCTGGCGGTATATGGAACCCCTCCCAAAAGCCCCAGGCGCAGGACTGCC 1045
Oy 1048 GACTCACTCTCCAGTGGAGCAGCA 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1046 CGACTCCCTCCCGGGGGAACCA 1070

```

RESULT 4
AL516503/c
LOCUS AL516503 LTI_NFL011.NBC1 962 bp mRNA linear EST 13-FEB-2001
DEFINITION AL516503 LTI_NFL011.NBC1 Homo sapiens cDNA clone CS0DA006YH04 3
PRIMER mRNA sequence.
ACCESSION AL516503
VERSION AL516503.1 GI:12779996
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
 source
 1..962
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DA006YH04"
 /clone_11b="LTI_NFL011.NBC1"
 /sex="male"
 /tissue.type="neuroblastoma cells"
 /lab_host="DH10B"
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-Oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library is not normalized, but is the control for
 the normalized libraries. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a

```

BASE COUNT      193 a      262 c      248 g      243 t      16 others
ORIGIN
Query Match      74.4% Score 884.2 DB 9: Length 962:
Best Local Similarity 95.8% Pred No. 6.8e-206:
Matches 920: Conservative 13: Mismatches 23: Indels 4: Gaps 3:

```

division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

```

Oy 232 CGGATCTCTGAGCCCTGGTTGAACATCTCATCCCTGTGTAGACCGGATCCGATATGTG 291
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 962 CGGATCTCTGAGCCCTGGTTGAACATCTCATCCCTGTGTAGACCGGATCCGATATGTG 904
Oy 292 CAGAGTCTAAGGAAATTTGTCAATCAAGTCCCTGAGCAAGTCCGCTGTGACTCCCAAT 351
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 903 CAGACTCTCAAGGAAATTTCAATCAAGTCCCTGAGCAAGTCCGCTGTGACTCCCAAT 844
Oy 352 GTAACCTGCAATATGATGAGGAGCTCTTACAGCGCATCATGAGACCTTACAAAGGCAAG 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 GTAACCTGCAATATGATGAGGAGCTCTTACAGCGCATCATGAGACCTTACAAAGGCAAG 784
Oy 412 TACGGTGTGAGAGACCCCTGATGATGCGCTACACCAGCTAAGTCAAAACAACATGAGATCA 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 TACGGTGTGAGAGACCCCTGATGATGCGCTACACCAGCTAAGTCAAAACAACATGAGATCA 724
Oy 472 GAGCTCGGCAAACTCTCTTTGAGCAAAAGTCTCCGGGAAACGGGAGTCCCTGAATGCGAC 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 GAGCTCGGCAAACTCTCTTTGAGCAAAAGTCTCCGGGAAACGGGAAATCCCTGAATGCGAC 664
Oy 532 ATTGTGATGCCATCAACCAAGCTGTGACTGTGGGGTATCGCGTCCCTGCTGTTATGAG 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 ATTGTGATGCCATCAACCAAGCTGTGACTGTGGGGTATCGCGTCCCTGCTGTTATGAG 604
Oy 592 ATCAAGGATATCATGTGCGCCACCCTGGGTGAAGAGTCTATGCAGATGAGGTGAGGCA 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 ATCAAGGATATCATGTGCGCCACCCTGGGTGAAGAGTCTATGCAGATGAGGTGAGGCA 544
Oy 652 GAGCGGCGGAAACGGGCGCCAGTCTTAGAGTGTGAGGGAGCCGAGAGTCCGCTCATAT 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 GAGCGGCGGAAACGGGCGCCAGTCTTAGAGTGTGAGGGAGCCGAGAGTCCGCTCATAT 484
Oy 712 GTGGCAGAGGGAAGAAAGAGCCCGCATCTGCGCTCCGAGCAAGAAAGCTGAAGCA 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 483 GTGGCAGAGGGAAGAAAGAGCCCGCATCTGCGCTCCGAGCAAGAAAGCTGAAGCA 424
Oy 772 ATPAATCAGGACAGAGAGAGGCGCAGTGCAG--TTCTGGCGAAGGCCAAGGCTPAAGCTGA 830
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 ATPAATCAGGACAGAGAGAGGCGCAGTGCAGTCTTCTGGCGAAGGCCAAGGCTPAAGCTGA 364
Oy 831 AGCTATTCGAATCTGCGCTGCGAGCTGTGACACAACATATGAGAGTGCAGCGCTTCACT 890
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 AGCTATTCGAATCTGCGCTGCGAGCTGTGACACAACATATGAGAGTGCAGCGCTTCTTACT 304
Oy 891 GACTGTGGCGAGAGTATGTGAGCGCGTCTTCAAACTGCGCAAGAGTCTCAACACTAT 950
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 GACTGTGGCGAGAGTATGTGAGCGCGCTTCTTCAAACTGCGCAAGAGTCTCAACACTAT 244
Oy 951 CCTACTGCGCTTCAACCTGCGAGTGTCAACAGCATGTGGCTCAGGCGCATGGGTGTATA 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 CCAACTGCGCTTCAACCTGCGAGTGTCAACAGCATGTGGCTCAGGCGCATGGGTGTATA 184
Oy 1011 TGGAGCCCTCACCAAAAGCCCGCAGTGGCAGGACTCCAGACTCACTCTCCAGTGGAGCAG 1070
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 TGGAGCCCTCACCAAAAGCCCGCAGTGGCAGGACTCCAAACTCACTCTCCAGTGGAGCAG 124
Oy 1071 CAGAGATGTCAGGGTACAGTGAAGTGTGATGAGAACTTGTATCGATCGATGAAGTGAAG 1130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 CAGAGATGTCAGGGTACAGTGAAGTGTGATGAGAACTTGTATCGATCGATGAAGTGAAG 64
Oy 1131 TTAGTGAAGCTGGGCTTNGCCAGGAGTCTGGGGACAAAG--AAGCAGATTTTCTGATTT 1188

```



```
DB 63 TTATGGAGCTGGGCTTGGCCAGGAGATYGGGACAAAGATATGCAGATTTCCTTT 4
|||||
RESULT 5
BM468527 994 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6475743 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578198
DEFINITION 5', mRNA sequence.
ACCESSION BM468527
VERSION BM468527.1 GI:18517569
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 994)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DMP
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12332 row: n column: 23
High quality sequence stop: 658.
Location/Qualifiers
1..994
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5578198"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 225 a 276 c 292 g 194 t 7 others
ORIGIN
Query Match 73.4%; Score 871.8; DB 10; Length 994;
Best Local Similarity 97.5%; Pred. No. 7.6e-203;
Matches 912; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 49 CTGCGGCGGGGGAATGCGGGCGGGCGGGGCGGACACGCGGCGCTTTTGTCTGAG 108
|||||
DB 1 CTGCGGCGGGGGAATGCGGGCGGGCGGGCGGCGGCGGCGGCGGCGGCGCTTTTGTCTGAG 59
|||||
QY 109 GGGCTCTACTGCTTCTGCGCGCGCTCCG-GCGGCGCTCTCTGATGCCCCGAAA 167
|||||
DB 60 GGGCTCTACTGCTTCTGCGCGCGCTCCGCGCGGCGCTCTCTGATGCCCCGAAA 119
|||||
QY 168 CACCGTGTACTGCTTCTGCGCGCGACGAGGCGCTGGGTGGAGGAAATGGCGGATT 227
|||||
DB 120 CACCGTGTACTGCTTCTGCGCGCGACGAGGCGCTGGGTGGAGGAAATGGCGGATT 179
|||||
QY 228 CCACCGGATCCGAGCGCTGTTGAAATCCATCCATCCCTGTATACCGGATCCGATA 287
|||||
DB 180 CCACCGGATCCGAGCGCTGTTGAAATCCATCCATCCCTGTATACCGGATCCGATA 239
|||||
QY 288 TTGTGAGAGTCTCAAGAAATTTGATCAACGTCGCTGAGCAGTGGCTGTGACTCGA 347
|||||
DB 240 TTGTGAGAGTCTCAAGAAATTTGATCAACGTCGCTGAGCAGTGGCTGTGACTCGA 299
|||||
QY 348 CAATGTACTCTGCAAAATGATGAGTCTTTTACTCTGCGATCATGACCTTACAAAGC 407
|||||
```

```
DB 300 CAATGTACTCTGCAAAATGATGAGTCTTTTACTCTGCGATCATGAGACCTTTACAAAGC 359
|||||
QY 408 AAGCTAGGTGTGAGAGACCTTGATATGCGCTACCCAGCTACTACTAAACACCATGAG 467
|||||
DB 360 AAGCTAGGTGTGAGAGACCTTGATATGCGCTACCCAGCTACTACTAAACACCATGAG 419
|||||
QY 468 ATCAGAGCTGGCAAACTCTCTNTGACAAAGTCTTCGGGAAAGGAGTCCCTGAATGC 527
|||||
DB 420 ATCAGAGCTGGCAAACTCTCTGACAAAGTCTTCGGGAAAGGAGTCCCTGAATGC 479
|||||
QY 528 CAGCATTTGTGATGTCATCAACCAAGCTGCTGACTGTGGGGATATCCGTGCTGCTTA 587
|||||
DB 480 CAGCATTTGTGATGTCATCAACCAAGCTGCTGACTGTGGGGATATCCGTGCTGCTTA 539
|||||
QY 588 TGAGATCAAGGATATCCATGTGCGACCCCGGGTGAAGAGTCTATGAGATGCAAGTGA 647
|||||
DB 540 TGAGATCAAGGATATCCATGTGCGACCCCGGGTGAAGAGTCTATGAGATGCAAGTGA 599
|||||
QY 648 GCGAGAGCGGCGGAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGATCGGCAT 707
|||||
DB 600 GCGAGAGCGGCGGAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGATCGGCAT 659
|||||
QY 708 CAATGTGGCAGAAAGGAAACAGGCCAGATCTTGCGCTCCGAGCAGAAAAGCTGA 767
|||||
DB 660 CAATGTGGCAGAAAGGAAACAGGCCAGATCTTGCGCTCCGAGCAGAAAAGCTGA 719
|||||
QY 768 ACAGATTAATCAGGACAGAGAGGCGGAGTCTGCGGAGGAGCCAAAGCTTAAGC 827
|||||
DB 720 ACAGATTAATCAGGACAGAGAGGCGGAGTCTGCGGAGGAGCCAAAGCTTAAGC 779
|||||
QY 828 TGAAGCTATTCGAATCCTGCTGAGCTGTGACACAAATATGAGATGACGAGCTTC 887
|||||
DB 780 TGAAGCTATTCGAATCCTGCTGAGCTGTGACACAAATATGAGATGACGAGCTTC 839
|||||
QY 888 ACTGACTGTGGCGGAGCAGATATGTCACGCGGTTCTCCAAATGCGCCAAAGCTCAAC 947
|||||
DB 840 ACTGACTGTGGCGGAGCAGATATGTCACGCGGTTCTCCAAATGCGCCAAAGCTCAAC 899
|||||
QY 948 TATCTACTGCTCCCTCCACCTGCGGATGTCACC 981
|||||
DB 900 TATCTACTGCTCCCTCCACCTGCGGATGTCACC 934
|||||

RESULT 6
BM468407 998 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6475436 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578133
DEFINITION 5', mRNA sequence.
ACCESSION BM468407
VERSION BM468407.1 GI:18517449
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DMP
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12332 row: 1 column: 06
High quality sequence stop: 656.
Location/Qualifiers
1..998
/organism="Homo sapiens"
source
```


OY	53	GGGGTGGGAAATGCTGGCCGGCGCGGGGGGACATGAGGGCCCTTTTGTGAGGGGC	112
Db	1	GGCGGTGGGAAATGCTGGCGCGCGCGGGG-GAGCGCACTGGGGGCCCTTTTGTGAGGGGC	59
OY	113	TCTCTACTGCGCTTCGCGCGCGCGCTCCG-GCGCGGCTCTCTGTGATGTCGCCGAAACAC	171
Db	60	TCTCTACTGCGCTTCGCGCGCGCGCTCCGCGCGCGCTCTCTGTGATGTCGCCGAAACAC	119
OY	172	GTGTACTGTTTCGTCCCGACAGCAGGAGCCCTGGGTGTGTGAGCCGATGGCCCGATTCCAC	231
Db	120	GTGTACTGTTTCGTCCCGACAGCAGGAGCCCTGGGTGTGTGAGCCGATGGCCCGATTCCAC	179
OY	232	CGGATCCTGGAGCCCTGGTTTGAACTCCGATCCCTGTGTAGACCGGATCGATATCTG	291
Db	180	CGGATCCTGGAGCCCTGGTTTGAACTCCGATCCCTGTGTAGACCGGATCGATATCTG	239
OY	292	CAGACTCTCAAGGAATTTGTCTATCAACGTCGCTGAGCAGTGGCGCTGTACTCTGCACAT	351
Db	240	CAGACTCTCAAGGAATTTGTCTATCAACGTCGCTGAGCAGTGGCGCTGTACTCTGCACAT	299
OY	352	GTAACCTCTGCAAAATCGATGAGATCCTTTTACGTGGGCATCATGAGACCTTACAGGCAAC	411
Db	300	GTAACCTCTGCAAAATCGATGAGATCCTTTTACGTGGGCATCATGAGACCTTACAGGCAAC	359
OY	412	TACGCTGTGAGAGACCCCTGAGTATCCGCTTCACCCAGCTAGCTCAAAACAACATGATGATCA	471
Db	360	TACGCTGTGAGAGACCCCTGAGTATCCGCTTCACCCAGCTAGCTCAAAACAACATGATGATCA	419
OY	472	GAGCTTGGCAAACTCTGTTNTGACAAAGTCTTTCGGGGAACGGGAGTCCCTGATGCCAGC	531
Db	420	GAGCTTGGCAAACTCTCTGTGTGACAAAGTCTTTCGGGGAACGGGAGTCCCTGATGCCAGC	479
OY	532	ATTGTGATGCCATCAACCAAGCTGCTGACGTGCTGGGATATCCGCTGCTGCTGATATAG	591
Db	480	ATTGTGATGCCATCAACCAAGCTGCTGACGTGCTGGGATATCCGCTGCTGCTGATATAG	539
OY	592	ATCAAGGATATCATGTGCCACCCCGGGTGAAGAAGTCTATGACAGTCCAGTGGAGGCA	651
Db	540	ATCAAGGATATCATGTGCCACCCCGGGTGAAGAAGTCTATGACAGTCCAGTGGAGGCA	599
OY	652	GAGCGGGGAAACGGGGCCACAGTTTCTAGAGTGTGAGGGAGACCCGAGTCCGCAATCAT	711
Db	600	GAGCGGGGAAACGGGGCCACAGTTTCTAGAGTGTGAGGGAGACCCGAGTCCGCAATCAT	659
OY	712	GTTGCGAAGAGGAAGAACAGAGCCACAGATCTCTGGCCCTCCGAAGCAGAAAGGCTTAAACAG	771
Db	660	GTTGCGAAGAGGAAGAACAGAGCCACAGATCTCTGGCCCTCCGAAGCAGAAAGGCTTAAACAG	719
OY	772	ATTAATTCAGGCGACGAGAGAGGCCAGTGACGTTCTTGGCGAAGGCCAAGGCTTAAAGCTGAA	831
Db	720	ATTAATTCAGGCGACGAGAGAGGCCAGTGACGTTCTTGGCGAAGGCCAAGGCTTAAAGCTGAA	779
OY	832	GCTATTTCGAATCTGTGGCTGACGCTTGACACACAACATTAATGGAGATGCAAGACTTCACTG	891
Db	780	GCTATTTCGAATCTGTGGCTGACGCTTGACACACAACATTAATGGAGATGCAAGACTTCACTG	839
OY	892	ACTGTGGCCG 901	
Db	840	ACTGTGGCCG 849	

RESULT	10
BM453159	
LOCUS	BM453159 1061 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOCURT_6387825 NIH_MGC_71 Homo sapiens cDNA IMAGE:5529785 5', mRNA sequence.
ACCESSION	BM453159
VERSION	BM453159.1 GI:18502199
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1061)	NIH-MGC	http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.	Email: cgapbs-remail.nih.gov	Tissue Procurement: ATCC	CDNA Library Preparation: Life Technologies, Inc.	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	http://image.llnl.gov	Plate: LLM12208	row: m column: 18
High quality sequence stop: 661.	Location/Qualifiers	1. 1061		
FEATURES	source	1. 1061		
BASE COUNT	259 a	298 c	279 g	225 t
ORIGIN				
Query Match	69.1%;	Score 821;	DB 10;	Length 1061;
Best Local Similarity	98.8%;	Pred. No. 2.2e-190;		
Matches 835;	Conservative	0;	Mismatches 9;	Indels 1;
Gaps				
QY	345	CGACAAATGTAACCTGCGCAAAATCGATGAGTGCCTTTACCTGCGCATATGAGACCTTACAA	404	
DB	1	CGACAAATGTAACCTGCGCAAAATCGATGAGTGCCTTTACCTGCGCATATGAGACCTTACAA	60	
QY	405	GCGAAGCTAGCGGTGTGGAGGACCCCTGATGTGCGCGTACCACCACTGATGCTCAAAACACAT	464	
DB	61	GCGAAGCTAGCGGTGTGGAGGACCCCTGATGTGCGCGTACCACCACTGATGCTCAAAACACAT	120	
QY	465	GAGATCAGAGCTCGGCAAACTCTCTTNGACAAGATCTTCCGGGAACGGGAGTCCCTGAA	524	
DB	121	GAGATCAGAGCTCGGCAAACTCTCTTNGACAAGATCTTCCGGGAACGGGAGTCCCTGAA	180	
QY	525	TGCCAGCATTTGTGATGATGCATCAACACAGCTGCTGACTGCTGGGGTATCCGCTGCTTNG	584	
DB	181	TGCCAGCATTTGTGATGATGCATCAACACAGCTGCTGACTGCTGGGGTATCCGCTGCTTNG	240	
QY	585	TTATGAGATCAAGATATTCATGTCATGTCACCCGGGGTGAAGATCTATTCAGATGTCAGGT	644	
DB	241	TTATGAGATCAAGATATTCATGTCATGTCACCCGGGGTGAAGATCTATTCAGATGTCAGGT	300	
QY	645	GGAGGCAAGAGCGGGGGAAGCGGGCCACAGTCTTAAGATCTGAAGGGGACCCGAGAGTCCGC	704	
DB	301	GGAGGCAAGAGCGGGGGAAGCGGGCCACAGTCTTAAGATCTGAAGGGGACCCGAGAGTCCGC	360	
QY	705	CATCAATTTGGCACAAGGGAAGAAACAGGCCACAGATCTGTGGCTCCGGAAGCAGAAAAAGC	764	
DB	361	CATCAATTTGGCACAAGGGAAGAAACAGGCCACAGATCTGTGGCTCCGGAAGCAGAAAAAGC	420	
QY	765	TGAACAGATTAATCAGGAGCAGAGAGAGGCGCAGTCTGCGGCAAGGCAAGGCTTAA	824	
DB	421	TGAACAGATTAATCAGGAGCAGAGAGAGGCGCAGTCTGCGGCAAGGCAAGGCTTAA	480	
QY	825	AGCTGAAGCTATTCGAATCTCTGGCTGCGAGCTCTGACACAACAATATGAGATGACAGAC	884	
DB	481	AGCTGAAGCTATTCGAATCTCTGGCTGCGAGCTCTGACACAACAATATGAGATGACAGAC	540	
QY	885	TTTCACTGACTGTGGCCGAGCAGTATGTCTACAGCGGCTTCTCCAACTGGCCAGAGACTCCAA	944	
DB	541	TTTCACTGACTGTGGCCGAGCAGTATGTCTACAGCGGCTTCTCCAACTGGCCAGAGACTCCAA	600	

QY 945 CACTATCTACTGCCCCCTCCACCCCTGGCGATGTCACACAGATGTGGCTCAGGCCATGGG 1004
 |||||
 DB 601 CACTATCTACTGCCCCCTCCACCCCTGGCGATGTCACACAGATGTGGCTCAGGCCATGGG 660
 QY 1005 TGTATATGAGACCCCTTCACCAAGCCCCAGTCCAGGGAGCTCCAGACTCTCTCCAGTGG 1064
 |||||
 DB 661 TGTATATGAGACCCCTTCACCAAGCCCCAGTCCAGGGAGCTCCAGACTCTCTCCAGTGG 720
 QY 1065 GAGCAGCAGAGATGTCCAGAGTACAGATCAAGTNTTGTATGAGAACTTGATCGAGTCAA 1124
 |||||
 DB 721 GAGCAGCAGAGATGTCCAGAGTACAGATCAAGTNTTGTATGAGAACTTGATCGAGTCAA 780
 QY 1125 GATGAGTATAGTACAGTGGCTGGCTTTCAGGAGAGTCT-6GGGCAAGAGAGAGAGATTTCG 1183
 |||||
 DB 781 GATGAGTATAGTACAGTGGCTGGCTTTCAGGAGAGAGTCTGGGGGCAAGAGAGATTTCG 840
 QY 1184 TGATT 1188
 |||||
 DB 841 TGATT 845

RESULT 11
 AL525167/c 899 bp mRNA linear EST 13-FEB-2001
 LOCUS DEFINITION AL525167 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC005YF20 3
 prime, mRNA sequence.
 ACCESSION AL525167
 VERSION AL525167.1 GI:12788660
 KEYWORDS EST.
 SOURCE human.
 ORGANISM human.

REFERENCE 1 (bases 1 to 899)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source location/Qualifiers
 1. 899

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC005YF20"
 /clone_11b="LTI_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 175 a 246 c 228 g 221 t 29 others
 ORIGIN

Query Match 67.6% Score 802.6 DB 9: Length 899;
 Best Local Similarity 94.2% Pred. No. 6.3e-186;
 Matches 800; Conservative 24; Mismatches 25; Indels 0; Gaps 0;

QY 340 ACTCTGCACATGATTAAGTTCGAAATCATGATGAGTCTTACTGGGCGATCGATGAGACCT 399
 |||||
 DB 899 ACTCTGCACATGATTAAGTTCGAAATCATGATGAGTCTTACTGGGCGATCGATGAGACCT 840

QY 400 TACAGGCAAGCTACGGTGTGAGAGACCCCTGATGATGCCGTACCCAGCTAGCTAACA 459
 :|||
 DB 839 HACAGGCAAGCTACGGTGTGAGAGACCCCTGATGATGCCGTACCCAGCTAGCTAACA 780
 QY 460 ACCATGAGATCAGAGCTCGGCAAACTCTTNTGGAAGAGTCTTCGGGAACGGGAGTCC 519
 |||||
 DB 779 ACCATGAGATCAGAGCTCGGCAAACTCTTNTGGAAGAGTCTTCGGGAACGGGAGTCC 720
 QY 520 CTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTGCTGGGATTCGGCTGC 579
 |||||
 DB 719 CTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTGCTGGGATTCGGCTGC 660
 QY 580 CTNCGTTATGATCAAGATATCCATGTGCCACCCCGGGTAAAGAGTCTATGAGATG 639
 |||||
 DB 659 CTNCGTTATGATCAAGATATCCATGTGCCACCCCGGGTAAAGAGTCTATGAGATG 600
 QY 640 CAGGTGAGGCGAGCGGGGGAACGGGCGCACAGTTCTAGACTGTGAGGGAGCCGAGAG 699
 |||||
 DB 599 CAGGTGAGGCGAGCGGGGGAACGGGCGCACAGTTCTAGACTGTGAGGGAGCCGAGAG 540
 QY 700 TCGGCCATCAATGTGGCAGAGGAAGAAACAGGCCAGATCTGGCCTCCGAGCAGAA 759
 |||||
 DB 539 TCGGCCATCAATGTGGCAGAGGAAGAAACAGGCCAGATCTGGCCTCCGAGCAGAA 480
 QY 760 AAGGCTGAACAGATTAATCAGCAGCAGAGAGGCCAGTGTGCGGAAGGCCAAG 819
 |||||
 DB 479 AAGGCTGAACAGATTAATCAGCAGCAGAGAGGCCAGTGTGCGGAAGGCCAAG 420
 QY 820 GCTAAAGCTGAAGCTATTCGATTCCTGCGCTGCAGCTGTACACATATATGAGATGCA 879
 |||||
 DB 419 GCTAAAGCTGAAGCTATTCGATTCCTGCGCTGCAGCTGTACACATATATGAGATGCA 360
 QY 880 GCAGCTTCACTACTGTGGCGCCAGCAGATGTGTACAGGCCCTTCCAAACTGGCCAGAGAC 939
 |||||
 DB 359 GCAGCTTCACTACTGTGGCGCCAGCAGATGTGTACAGGCCCTTCCAAACTGGCCAGAGAC 300
 QY 940 TCCAACTACTATCTTACTGCGCTTCACACCTGGCGATGTACACAGATGTGCTCAGGCC 999
 |||||
 DB 299 TCCAACTACTATCTTACTGCGCTTCACACCTGGCGATGTACACAGATGTGCTCAGGCC 240
 QY 1000 ATGGGCTATATGAGAGCCCTCCACCAAGCCCAAGGCCAGGAGCTCCAACTGACTCTCC 1059
 |||||
 DB 239 ATGGGCTATATGAGAGCCCTCCACCAAGCCCAAGGCCAGGAGCTCCAACTGACTCTCC 180
 QY 1060 AGTGGAGCAGCAGAGATGTCCAGAGTACAGATGCAAGTNTTGTATGAGAGACTTATGCA 1119
 |||||
 DB 179 AGTGGAGCAGCAGAGATGTCCAGAGTACAGATGCAAGTNTTGTATGAGAGACTTATGCA 120
 QY 1120 GTCAAGATGATTTAGTGTGAGCTGGCTTGGCCAGGAGTCTGGGCAAGAGACAGATT 1179
 |||||
 DB 119 GTCAAGATGATTTAGTGTGAGCTGGCTTGGCCAGGAGTCTGGGCAAGAGACAGATT 60
 QY 1180 TTCCGTGATT 1188
 |||||
 DB 59 TTCCGTGATT 51

RESULT 12
 BM423347 1095 bp mRNA linear EST 29-JAN-2002
 LOCUS DEFINITION BM423347
 AGENCOUFT_6402181 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5516467
 5', mRNA sequence.

ACCESSION BM423347
 VERSION BM423347.1 GI:18391559
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1095)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLCM2017 row: b column: 20
High quality sequence stop: 503.

FEATURES
source
1. 1095
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5516467"
/clone_lib="NIH_MGC_41"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 285 a 282 c 302 g 226 t

ORIGIN

Query Match 66.7%; Score 792.6; DB 10; Length 1095;
Best Local Similarity 96.5%; Pred. No. 2e-183;
Matches 831; Conservative 0; Mismatches 26; Indels 4; Gaps 2;

QY 203 GGGTGGTGGAGCGCAATGGCCGATTCACCGATCTGGAGCCTGTTGAACATCTCTCA 262
DB 1 GGGTGGTGGAGCGCAATGGCCGATTCACCGATCTGGAGCCTGTTGAACATCTCTCA 60
QY 263 TCCCTGTGTAGACCGGATCCGATGTGTGAGAGTCTCAAGGAATGTGATCAACGCTGC 322
DB 61 TCCCTGTGTAGACCGGATCCGATGTGTGAGAGTCTCAAGGAATGTGATCAACGCTGC 120
QY 323 CTGAGCAGTCGGCTGTGACTCTGACAAATGTAACTGTGCAATGTGATGAGTCTTACC 382
DB 121 CTGAGCAGTCGGCTGTGACTCTGACAAATGTAACTGTGCAATGTGATGAGTCTTACC 180
QY 383 TGGGCATCATGACCCCTTACAGGCAAGCTAGCGTGTGAGAGACCTGAGTATGCGGTCA 442
DB 181 TGGGCATCATGACCCCTTACAGGCAAGCTAGCGTGTGAGAGACCTGAGTATGCGGTCA 240
QY 443 CCAGCTAGCTCAAAACACCATGAGATCAGAGTCGGCAAACTCTGNTGACAAATCTT 502
DB 241 CCAGCTAGCTCAAAACACCATGAGATCAGAGTCGGCAAACTCTGNTGACAAATCTT 300
QY 503 TCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTGACT 562
DB 301 TCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTGACT 360
QY 563 GCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCGGGTGA 622
DB 361 GCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCGGGTGA 420
QY 623 AAGAGCTATGATGATGAGGTGAGGAGGCGGCGGAAACGGCGCACAGTTCTAGAGT 682
DB 421 AAGAGCTATGATGATGAGGTGAGGAGGCGGCGGAAACGGCGCACAGTTCTAGAGT 480
QY 683 CTGAGGGAGACCCGAGAGTGGCCATCAATGTGACAGAAAGGAGAAACAGGCCAGATCC 742
DB 481 CTGAGGGAGACCCGAGAGTGGCCATCAATGTGACAGAAAGGAGAAACAGGCCAGATCC 540
QY 743 TGGCCTCCGAGAGAGAAAGGCTGAACAGTAATCAGGACGACGAGAGCCAGTGCAG 802

DB 541 TGGCCTCCGAGAGCAAGAAAGGCTGACAGATTAATTCAGCAGCAGAGAGCCAGTGCAG 600
QY 803 TTTCGGGCAAGGCCAAGGCTAAAGCTGAGCTATTTGCAATFCTGGCTGACCTGACAC 862
DB 601 TTTCGGGCAAGGCCAAGGCTAAAGCTGAGCTATTTGCAATFCTGGCTGACCTGACAC 660
QY 863 AACATAATGAGATGACAGCAGCTTACACTGTGTGGCGAGCAGATGATGAGCGGTTCT 922
DB 661 AACATAATGAGATGACAGCAGCTTACACTGTGTGGCGAGCAGATGATGATGAGCGGTTCT 720
QY 923 CCAAACTGGCCCAAGAGACTCCAACTATCTACTGCCCCCTCCACCCCTGGCGATGCACCA 982
DB 721 CCAAACTGGCCCAAGAGACTCCAACTATCTACTGCCCCCTCCACCCCTGGCGATGCACCA 780
QY 983 GCATGCTGCTGAGGCTCAAGCCTGATGATGAGAGCCCTCCACCCCGCTGCC 1038
DB 781 GCATGCTGCTGAGGCTCAAGCCTGATGATGAGAGCCCTCCACCCCGCTGCC 840
QY 1039 GGCAGCTCCAGACTCACTCTCC 1059
DB 841 GGCAGCTCCAGACTCACTCTCC 861

RESULT 13
BM459766 918 bp mRNA linear EST 05-FEB-2002
LOCUS BM459766
DEFINITION AGENCOURT_6417777 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:5534403
ACCESSION BM459766
VERSION BM459766.1 GI:18508806
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLCM2220 row: n column: 04
High quality sequence stop: 626.

FEATURES
source
1. 918
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5534403"
/clone_lib="NIH_MGC_71"
/tissue_type="leptomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-Sport6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

BASE COUNT 227 a 249 c 264 g 175 t 3 others

ORIGIN

Query Match 66.6%; Score 791.8; DB 10; Length 918;
Best Local Similarity 97.2%; Pred. No. 2.8e-183;
Matches 836; Conservative 0; Mismatches 17; Indels 7; Gaps 3;

QY 179 TGTTCGTGGCGGACAGAGGCTGGGTGTGAGCGCAATGGCCGATTCACCGGATCC 238
DB 13 TGTTCGTGGCGGACAGAGGCTGGGTGTGAGCGCAATGGCCGATTCACCGGATCC 72

OY	239	TGGAGCCCTGTTTGAACATCCGCATCCCTGGTTTAGAGCGGATCCGATTGTGGAGATC	298
Db	73	TGGAGCCCTGTTTGAACATCCGCATCCCTGGTTTAGAGCGGATCCGATTGTGGAGATC	132
OY	299	TCAAGGAATTTGTCAATCAAGCTGCCGTGAGCAGTCGGCTGTACTCTCGACAATGTAACTC	358
Db	133	TCAAGGAATTTGTCAATCAAGCTGCCGTGAGCAGTCGGCTGTACTCTCGACAATGTAACTC	192
OY	359	TGCAATTCGATGAGTACTCTTTAACCCTTGCGCATCATGAGACCTTTACAAGCAAGCTAACGTG	418
Db	193	TGCAATTCGATGAGTACTCTTTAACCCTTGCGCATCATGAGACCTTTACAAGCAAGCTAACGTG	252
OY	419	TGGAGGACCCGTGAGTATCCCGTCAACCCAGGTAAAGCTCAACAACCATGTAGATCAGAAGTCG	478
Db	253	TGGAGGACCCGTGAGTATCCCGTCAACCCAGGTAAAGCTCAACAACCATGTAGATCAGAAGTCG	312
OY	479	GCAAACTCTCTNTGAGACAAGTCTTCCGGGAACGGAGATCCCTGTAATCCAGCATTTGGG	538
Db	313	GCAAACTCTCTNTGAGACAAGTCTTCCGGGAACGGAGATCCCTGTAATCCAGCATTTGGG	372
OY	539	ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCCTNCGTTATGAGATCAAG	598
Db	373	ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCCTNCGTTATGAGATCAAG	432
OY	599	ATATTCATGTGTCACACCCCGGGTGAAGAGTCTATGACATGACAGTGGAGGACAGCCGC	658
Db	433	ATATTCATGTGTCACACCCCGGGTGAAGAGTCTATGACATGACAGTGGAGGACAGCCGC	492
OY	659	GGAACGGGGCCACAGTTCGTAGAGTGTGAGGGGAGCCGAGAGTCCGGCATCATATGTGGCAG	718
Db	493	GGAACGGGGCCACAGTTCGTAGAGTGTGAGGGGAGCCGAGAGTCCGGCATCATATGTGGCAG	552
OY	719	AAGGGAAGAAACAGGCCCCAGATCTCGGCCCTCGAAGCAGAAAAGGCTGACAGATAATC	778
Db	553	AAGGGAAGAAACAGGCCCCAGATCTCGGCCCTCGAAGCAGAAAAGGCTGACAGATAATC	612
OY	779	AGGCAGCAGGAGAGGCCACAGTTCGATTTCTGGGGAAGGCCAAGGCTTAAAGCTGAAGCTATTC	838
Db	613	AGGCAGCAGGAGAGGCCACAGTTCGATTTCTGGGGAAGGCCAAGGCTTAAAGCTGAAGCTATTC	672
OY	839	GAATCTCGGCTGCACGCTGTGACACAAACATTAATGAGATGACAGCAGCTTCACGTGACTGTGG	898
Db	673	GAATCTCGGCTGCACGCTGTGACACAAACATTAATGAGATGACAGCAGCTTCACGTGACTGTGG	732
OY	899	CCGAGCAGTATGTCAGCGCGCTTCTCCAACATGCGCCAAAGAGCTCCAACACTATCTACTGC	958
Db	733	NCGAGCAGTATGTCAGCGCGCTTCTCCAACATGCGCCAAAGAGCTCCAACACTATCTACTGC	792
OY	959	CCTCCACACCCCTGG-CGATGTCAACAGCATGTGTGGCTCACG---CCATGGGTGTATATGA	1014
Db	793	CCTCCACACCCCTGGCGGATGTCAACAGCATGTGGGGGCTTAAGGCCCATGTGGGTGTATATGN	852
OY	1015	---GGCCTCACCAAAGCCCC 1031 	
Db	853	AGCCCCTCACCAAAGCCCC 872	
RESULT	14		
LOCUS	BGA79080	943 bp	mRNA linear EST 21-MAR-2001
DEFINITION	605526159t1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4649846 5'		
ACCESSION	BGA79080		
VERSION	BGA79080.1	GI:13411359	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		

COMMENT

Contact: Robert Strausberg Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://Image.lnl.gov>
Plate: L1CMI430 row: e column: 15
High quality sequence stop: 857.

FEATURES

Location/Qualifiers

```

1..943
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4649846"
    /clone_1id="NIH_MGC_21"
    /tissue_type="choriocarcinoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: placenta; Vector: pOTR1; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >500bp for average insert size 1 kb." library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

```

BASE COUNT

	210 a	254 c	290 g	189 t
ORIGIN				

Query Match 66.6% Score 791.4 DB 10 Length 943;
Best Local Similarity 98.4% Pred. No. 3.6e-183;

Matches 861; Conservative 0; Mismatches 8; Indels 6; Gaps 6;

```

QY      32 GTTTGTTCGGAGGCTCGCTGGCGGCGTGGAATCTCTGGCCGCCGCGGGCGGGGCACT 91
        |||||||
DB       2 GTTTGTTCGGAGGCTCGCTGGCGGCGTGGAATCTCTGGCGCGGCGGCGGCACT 60
        |||||||
QY      92 GGGGCGCTTTTGCTGAGGGGCTCTACTGGCTTGCGCGCGCTCG -GCCGGGCTTC 150
        |||||||
DB       61 GGGGCGCTTTTGCTGAGGGGCTCTACTGGCTTGCGCGCGGCTCGCGGCGGCTTC 120
        |||||||
QY      151 TCTGATTGCCCCGGAACACCGTGCTCTGTCGTCGCCGACGAGAGGCTTGGGTGTG 210
        |||||||
DB       121 TCTGATTGCCCCGGAACACCGTGCTCTGTCGTCGCCGACGAGAGGCTTGGGTGTG 180
        |||||||
QY      211 GAGCGAATGGCGCGCATCCACCGGATCCCTGGAGCCTGCTTGAACATCTCATCCCTGTG 270
        |||||||
DB       181 GAGCGAATGGCGCGCATCCACCGGATCCCTGGAGCCTGCTTGAACATCTCATCCCTGTG 240
        |||||||
QY      271 TTAGACCGGATCCGATATGTGCAGAGCTTCAAAGAATAATTGCATCAACGTCGCTGAGCAG 330
        |||||||
DB       241 TTAGACCGGATCCGATATGTGCAGAGCTTCAAAGAATAATTGCATCAACGTCGCTGAGCAG 300
        |||||||
QY      331 TCGGCTGTGACTCTCGACAATGTAACTCTGCACAAATGCATGAGATCTTTACCTGCGCATC 390
        |||||||
DB       301 TCGGCTGTGACTCTCGACAATGTAACTCTGCACAAATGCATGAGATCTTTACCTGCGCATC 360
        |||||||
QY      391 ATGAGACCTTACAAAGGCAAGCTACGGTGTGAGAGACCCTGAGTATGCCCCTCACCCAGCTA 450
        |||||||
DB       361 ATGAGACCTTACAAAGGCAAGCTACGGTGTGAGAGACCCTGAGTATGCCCCTCACCCAGCTA 420
        |||||||
QY      451 GGTCAACAACCATAGATGACAGCTCGGCAAACTCTCTMNNGACAAAGTCTTTCGCGGAA 510
        |||||||
DB       421 GGTCAACAACCATAGATGACAGCTCGGCAAACTCTCTCTGAGCAAAAGTCTTTCGCGGAA 480
        |||||||
QY      511 CGGAGATCCCTGAATAGCCAGCAATTGTGATGATGCATCAACCAAGCTGTGACTGTGGGGT 570
        |||||||
DB       481 CGGAGATCCCTGAATAGCCAGCAATTGTGATGATGCATCAACCAAGCTGTGACTGTGGGGT 540
        |||||||
QY      571 ATCCGCTGCTTCGTTATGAGATCAAGATATCATGTGACCACCCGGGTGAAGAGTCT 630
        |||||||

```


Db 541 ATCCGCTGCCCTCCGTTATGATCAAGATATCCATGTGCCACCCGGGTGAAGAGCTCT 600
QY 631 ATGCAGATCGAGTGGAGGACGAGCGGGAACGGGCGACAGTTCTAGAGTCTGAGGG 690
Db 601 ATTCACATCGAGTGGAGGACGAGCGGGAACGGGCGACAGTTCTAGAGTCTGAGGG 660
QY 691 ACCCGAGTCCGGCCATCAATGTGGCAGGAAGGAAGAACAGGCCAGATCTGGGCTCC 750
Db 661 A-CCGAGAGTCGGCCATCAATGTGGCAGGAAGGAAGAACAGGCCAGATCTGGGCTCC 719
QY 751 GAAGCAAAAGGCTGAACAGATTAATCAAGCAGCAGAGAACGGCCAGTTCCT-GGC 809
Db 720 GAAGCGAAAGGCTGAACAGATTAATCAAGCAGCAGAGAACGGCCAGTTCCTGGC 779
QY 810 GAAGCGAAAGGCTGAACAGTGAAGCTATTCGATCTCTGGTGA-GCCTGACACACATA 868
Db 780 GAAGCGCGAGGTTAAAGCTGAAGCTATTCGATCTCTGGTGAAGCTTCTGACACATA 839
QY 869 ATGGAGATGACGAGCTTCTACTGCTGTGGCGAG 903
Db 840 ATGGAGATGACGAG--TTCACCTGAATGTGGCGAG 873
RESULT 15
BI862105 920 bp mRNA linear EST 10-OCT-2001
LOCUS 60339095F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5400023 5',
DEFINITION mRNA sequence.
ACCESSION BI862105
VERSION BI862105.1 GI:16002852
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 920)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgraphs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Place: LHAM12019 row: n column: 24
High quality sequence stop: 918.
Location/Qualifiers
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5400023"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT5; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 241 a 241 c 259 g 179 t
ORIGIN

Query Match 66.6%; Score 791.2; DB 10; Length 920;

Best Local Similarity 99.0%; Pred. No. 4e-183;

Matches 804; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 373 GTTCCTTACCTGGCATCATGACCTTACAAAGCAAGCTACGCTGTGAGAGACCTGTAG 432
|||||

Db 1 GTTCCTTACCTGGCATCATGACCTTACAAAGCAAGCTACGCTGTGAGAGACCTGTAG 60
QY 433 TATGCCCTCACACAGCTAGCTCAACAACATGATTCAGAGCTCGGCAAACTCTCTNTG 492
Db 61 TATGCCCTCACACAGCTAGCTCAACAACATGATTCAGAGCTCGGCAAACTCTCTCTG 120
QY 493 GACAAAGCTCTCCGGGAACGGAGATCCCTGAATCCAGCATTTGGATGATCAACCA 552
Db 121 GACAAAGCTCTCCGGGAACGGAGATCCCTGAATCCAGCATTTGGATGATCAACCA 180
QY 553 GCTGCTACTGCTGGGTATTCCTGCTTCGCTTATGATGATCAAGATATTCATGTCGA 612
Db 181 GCTGCTACTGCTGGGTATTCCTGCTTCGCTTATGATGATCAAGATATTCATGTCGA 240
QY 613 CCCCAGTGAAGAGCTATGCAATGATGAGTGAGGACGAGCGGCGAAACGGGCCACA 672
Db 241 CCCCAGTGAAGAGCTATGCAATGATGAGTGAGGACGAGCGGCGAAACGGGCCACA 300
QY 673 GTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCAATGATGAGGAGGAGAAACAG 732
Db 301 GTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCAATGATGAGGAGGAGAAACAG 360
QY 733 GCCAGATCTCTGGCTCCGAAGCAGAAAGGCTGAACAGATTAATCAAGCAGCAGAGAG 792
Db 361 GCCAGATCTCTGGCTCCGAAGCAGAAAGGCTGAACAGATTAATCAAGCAGCAGAGAG 420
QY 793 GCCAGTCAAGTCTGGGGAAGGCGCAAGGCTGAACGCTGAAGCTATTCGATCTGGCTGCA 852
Db 421 GCCAGTCAAGTCTGGGGAAGGCGCAAGGCTGAACGCTGAAGCTATTCGATCTGGCTGCA 480
QY 853 GCTCTGACACACATTAATGAGATGACAGCAGCTTCACTGACTGTGG--CCGAGCAGTATGT 911
Db 481 GCTCTGACACACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCCGAGAGTATGT 540
QY 912 CAGCGCTTCTCCAAACTGGCCAAAGACTCCACATATCTACTGCCCTCAACCTGG 971
Db 541 CAGCGCTTCTCCAAACTGGCCAAAGACTCCACATATCTACTGCCCTCAACCTGG. 600
QY 972 CGATGTCACAGCATGTGTGGCTCAGGCGCATGGGTTATGAGAGCCCTCACCAAGGCC 1031
Db 601 CGATGTCACAGCATGTGTGGCTCAGGCGCATGGGTTATGAGAGCCCTCACCAAGGCC 660
QY 1032 AGTGCAGGAGCTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGA 1091
Db 661 AGTGCAGGAGCTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGA 720
QY 1092 TGCAGATTTGATGAGGAGAACTTGATCGAGTCAAGATGAGTGTAGTGGAGCTTGGCC 1151
Db 721 TGCAGATTTGATGAGGAGAACTTGATCGAGTCAAGATGAGTGTAGTGGAGCTTGGCC 780
QY 1152 AGGAGATCTGGGGCAGCAAGCAAGCATTTTCC 1183
Db 781 AGGAGATCTGGGGCAGCAAGCAAGCATTTTTC 812

Search completed: September 22, 2002, 14:49:21
Job time: 5287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:32:18 : Search time 14.69 Seconds

(without alignments)
938.336 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLARAARGHMGPFAGELSTG.....RDVQGTDAKXDELDKRYKMS 356

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.5	160	VP12_BPPRD	P17637 bacterioph
2	8	2.2	343	UL14_HCMVA	P16756 human cytom
3	7	2.0	92	SV05_RAT	P50231 rattus norv
4	7	2.0	134	ATPE_RHORI	P05442 rhodospirill
5	7	2.0	158	RLZ9_HUMAN	P47914 homo sapien
6	7	2.0	164	RPOE_BACHD	O9K6D6 bacillus ha
7	7	2.0	166	KAC5_BOVIN	P04262 bos taurus
8	7	2.0	180	YFAZ_ECOLI	P76471 escherichia
9	7	2.0	194	UREF_HAEIN	P44395 haemophilus
10	7	2.0	228	LOLD_VIBCH	P57066 vibrio chol
11	7	2.0	232	RADC_BACHD	O9K8H4 bacillus ha
12	7	2.0	279	FIXA_RHILP	O05559 rhizobium l
13	7	2.0	281	REPL_BPP1	P19654 bacterioph
14	7	2.0	307	OTC_ARCFU	O29013 archaeoglob
15	7	2.0	344	IN37_SPIOL	P23325 spinacia ol
16	7	2.0	347	RECA_CHRVI	O32377 chromatium
17	7	2.0	372	TOLA_HAEIN	P44678 haemophilus
18	7	2.0	381	DUS6_HUMAN	O16828 homo sapien
19	7	2.0	381	DUS6_MOUSE	O9d6b1 mus musculu
20	7	2.0	381	DUS6_RAT	O64346 rattus norv
21	7	2.0	381	YE88_MYCTU	P71768 mycobacteri
22	7	2.0	421	TOLA_ECOLI	P19934 escherichia
23	7	2.0	422	KICW_HUMAN	O9C075 homo sapien
24	7	2.0	426	FLT1_DROME	O61491 drosophila
25	7	2.0	432	TTC_ECOLI	P22357 escherichia
26	7	2.0	433	Z001_YEAST	P32527 saccharomyc
27	7	2.0	435	MDHP_SPIOL	P52426 spinacia ol
28	7	2.0	448	TBA4_HUMAN	P05215 homo sapien
29	7	2.0	475	Z131_HUMAN	P52739 homo sapien
30	7	2.0	478	PKKH_ARATH	P50318 arabidopsis
31	7	2.0	480	PKKH_WHEAT	P12782 triticum ae
32	7	2.0	517	UZAF_SCHPO	P36629 schizosacch
33	7	2.0	517	YC09_CAMJE	O9pn86 campylobact

34	7	2.0	553	1	FXC1_HUMAN	Q12948 homo sapien
35	7	2.0	553	1	FXC1_MOUSE	O61572 mus musculu
36	7	2.0	570	1	IEP7_CAELA	P30642 caenorhabdi
37	7	2.0	580	1	RGP1_XENLA	O13066 xenopus lae
38	7	2.0	587	1	RGP1_HUMAN	P46061 homo sapien
39	7	2.0	589	1	RGP1_MOUSE	O9WU39 mus musculu
40	7	2.0	655	1	SCAG_MOUSE	P46062 mus musculu
41	7	2.0	693	1	SPAL_MOUSE	O95196 macaca mula
42	7	2.0	704	1	HS90_SCHRO	O54089 streptococc
43	7	2.0	706	1	SEM2_MACMU	P37937 schizophy11
44	7	2.0	739	1	RELA_STRDO	
45	7	2.0	891	1	MAZ3_SCHCO	

ALIGNMENTS

```

RESULT 1
ID      VP12_BPPRD      STANDARD:      PRT:      160 AA.
AC      P17637;
DT      01-AUG-1990 (rel. 15, Created)
DT      01-AUG-1990 (rel. 15, Last sequence update)
DT      01-DEC-1992 (rel. 24, Last annotation update)
DE      Single-stranded DNA-binding protein P12.
GN      XII.
OS      Bacteriophage PRD1.
OC      Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
OX      NCBI_TaxID=10658;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90152379; PubMed=2695404;
RA      Pakula T.M., Savilabti H., Bamford D.H.;
RT      "The organization of the right-end early region of bacteriophage PRD1
RT      genome."
RL      Gene 85:53-58(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90202706; PubMed=2180910;
RA      Gerendasy D., Ito J.;
RT      "Nucleotide sequence and transcription of the right early region of
RT      bacteriophage PRD1."
RL      J. Bacteriol. 172:1889-1898(1990).
CC      -I- FUNCTION: P12 HAS A REGULATORY EFFECT ON PHAGE DNA METABOLISM
CC      AND TRANSCRIPTION OF EARLY GENES. BINDS TO SINGLE-STRANDED DNA.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: M69077; AAA32476.1; -
DR      EMBL: M33478; AAA32447.1; -
DR      EMBL: M30146; AAA32453.1; -
DR      PIR: J00186; J00186.
DR      PIR: A35148; A35148.
KW      Early protein; DNA-binding.
SQ      SEQUENCE 160 AA; 16650 MW; 461004D25FC9F470 CRC64;

```

Query Match 2.5%; Score 9; DB 1; Length 160;

Best local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKAA 257

DB 152 AKAKAKAA 160

RESULT 2

```

UL14_HCMVA          STANDARD:          PRT:          343 AA.
ID UL14_HCMVA
AC P16756;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein UL14.
GN UL14.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Hornell T., Hutchinson C.A. III, Kourilides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17403; CA35447.1; -.
DR PIR: S09777; S09777.
DR HSSP: P04002; IATF.
KW Hypothetical protein.
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 343 AA: 38566 MW: F4A181A1AFDA1A98 CRC64;

Query Match          2.2%; Score 8; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 AAGEASAV 247
   |||||
Db 323 AAGEASAV 330

RESULT 3
SY05_RAT          STANDARD:          PRT:          92 AA.
ID SY05_RAT
AC P50231;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Small inducible cytokine A5 precursor (T-cell specific RANTES protein)
DE (SIS-delta).
GN SCYA5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=LONG EVANS; TISSUE=Lung;
RA Jones M.L., Shanley T.P., Ward P.A.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CHEMOKINE-ATTRACTANT FOR BLOOD MONOCYTES, MEMORY T HELPER
CC CELLS AND EOSINOPHILS. CAUSES THE RELEASE OF HISTAMINE FROM
CC BASOPHILS AND ACTIVATES EOSINOPHILS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06436; AAA96499.1; -.
DR HSSP: P13501; IRTN.
DR InterPro: IPR001811; Chemokine_IL8.
DR InterPro: IPR000827; Small_cytokine_CC.
DR Pfam: PF00048; IL8; 1.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; 1.
KW Cytokine; Chemotaxis; T-cell; Signal; Inflammatory response.
FT SIGNAL 1 24
FT CHAIN 25 92
FT DISULFD 34 58 SMALL INDUCIBLE CYTOKINE A5.
FT DISULFD 35 74 BY SIMILARITY.
SQ SEQUENCE 92 AA: 10170 MW: B4FEC2B4208ABC6 CRC64;

OY 272 AAASLTIV 278
   |||||
Db 5 AAASLTIV 11

Query Match          2.0%; Score 7; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 272 AAASLTIV 278
   |||||
Db 5 AAASLTIV 11

RESULT 4
ATPE_RHORU          STANDARD:          PRT:          134 AA.
ID ATPE_RHORU
AC P05442;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ATP synthase epsilon chain (EC 3.6.3.14) (ATP synthase F1 sector
DE epsilon subunit).
GN ATPC.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=85251588; PubMed=2861810;
RA Falk G., Hampe A., Walker J.E.;
RT "Nucleotide sequence of the Rhodospirillum rubrum atp operon.";
RL Biochem. J. 228:391-407(1985).
CC -1- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC H(+)(Out).
CC -1- SUBUNIT: F-type atpases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: A, B and C.
CC -1- SIMILARITY: BELONGS TO THE ATPASE EPSILON CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02499; CAB97256.1; -.
DR PIR: S08584; PMOFE.

```

DR HSP: P00832; IBSN.
 DR InterPro: IPR001469; ATP-synt_DE.
 DR Pfam: PF00401; ATP-synt_DE.1.
 DR Pfam: PF02823; ATP-synt_DE.N: 1.
 DR ProDom: PD000944; ATP-synt_DE.1.
 DR Hydrolase: ATP synthetase; Cf(1); Hydrogen ion transport.
 KW SEQUENCE 134 AA; 14307 MW; 8905084DF31C9222 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 134;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 SEAEKAE 235
 1111111
 DB 111 SEAEKAE 117

RESULT 5

RL29_HUMAN STANDARD; PRT; 158 AA.
 AC P47914:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 60S ribosomal protein L29 (Cell surface heparin binding protein HRP).
 GN RPL29.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RA LAW P.T.W., Tsui S.K.W., Lee C., Waye M.M.Y.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA MEDLINE=96180309; PubMed=8597591;
 RA Law P.T., Tsui S.K., Lam W.Y., Luk S.C., Hwang D.M., Liew C.C.,
 RA Lee C.Y., Fung K.P., Waye M.M.;
 RL "A novel cDNA encoding a human homologue of ribosomal protein L29";
 RL Biochim. Biophys. Acta 1305:105-108(1996).
 RN [3]
 RA SEQUENCE FROM N.A.
 RX MEDLINE=96216099; PubMed=8662616;
 RA Liu S., Smith S.E., Julian J., Rohde L.H., Karin N.J., Carson D.D.;
 RT "cDNA cloning and expression of HRP, a novel cell surface heparan
 sulfate/heparin-binding protein of human uterine epithelial cells and
 cell lines.";
 RL J. Biol. Chem. 271:11817-11823(1996).
 CC -1- SIMILARITY: BELONGS TO THE L29E FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U10248; AAC50499.1; -
 DR EMBL: 249148; CAAB9008.1; -
 DR EMBL: 049083; AAC50647.1; -
 DR MIM: 601832; -
 DR InterPro: IPR002673; Ribosomal_L29e.
 DR Pfam: PF01779; Ribosomal_L29e; 1.
 DR ProDom: PD010314; Ribosomal_L29e; 1.
 KW Ribosomal protein; Repeat; Heparin-binding.
 FT INIT MET 0 R -> A (TN REF. 2)
 FT CONFLICT 119 119 BY SIMILARITY.
 SQ SEQUENCE 158 AA; 17621 MW; 8FCC38682E6DDF28 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 158;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
 1111111
 DB 122 AKAKAKA 128

RESULT 6

RPOE_BACHD STANDARD; PRT; 164 AA.
 ID RPOE_BACHD
 AC O9K6D6:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable DNA-directed RNA polymerase delta subunit (RNAP delta
 factor).
 GN RPOE OR BH3793.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RA SEQUENCE FROM N.A.
 RX STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- FUNCTION: PARTICIPATES IN BOTH THE INITIATION AND RECYCLING PHASES
 CC OF TRANSCRIPTION. IN THE PRESENCE OF THE DELTA SUBUNIT, RNAP
 CC DISPLAYS AN INCREASED SPECIFICITY OF TRANSCRIPTION, A DECREASED
 CC AFFINITY FOR NUCLEIC ACIDS, AND AN INCREASED EFFICIENCY OF RNA
 CC SYNTHESIS BECAUSE OF ENHANCED RECYCLING (BY SIMILARITY).
 CC -1- SUBUNIT: RNAP IS COMPOSED OF A CORE OF 2 ALPHA, A BETA AND A
 CC BETA' SUBUNITS. THE CORE IS ASSOCIATED WITH A DELTA SUBUNIT AND
 CC ONE OF SEVERAL SIGMA FACTORS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RPOE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP001520; BAB07512.1; -
 DR DNA-directed RNA polymerase; Transcription: Complete proteome.
 KW DOMAIN 104 ASP/GLU-RICH (ACIDIC).
 FT SEQUENCE 164 AA; 19213 MW; 26E91F66DE67CA CRC64;

Query Match 2.0%; Score 7; DB 1; Length 164;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 250 KAKAKAE 256
 1111111
 DB 98 KAKAKAE 104

RESULT 7
 K2C5_BOVIN STANDARD; PRT; 166 AA.
 ID K2C5_BOVIN
 AC P04262:
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Keratin, type II cytoskeletal 68 kDa, component IB (Fragment).
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85128114; PubMed=6084625;
RA Jorcano J.L., Franz J.K., Franke W.M.;
RT "Amino acid sequence diversity between bovine epidermal cytokeratin
RT polypeptides of the basic (type II) subfamily as determined from cDNA
RT clones."
RL Differentiation 28:155-163(1984).
CC -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K03534; AAA30601.1; -.
DR PIR: A02948; KR802B.
DR HSSP: P10968; IMCG.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
DR KW Intermediate filament; Coiled coil; Keratin.
FT NON_TER 1 1
FT DOMAIN <1 41 ROD.
FT DOMAIN 42 166 TAIL.
FT DOMAIN <1 41 COIL_2B.
FT SEQUENCE 166 AA; 15505 MW; 4BFE495A7C6B4BE0 CRC64;

```

```

RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000315; AAC75310.1; ALT_INIT.
DR Ecogene: BG14088; yf4z.
DR Hypothetical protein; Signal; Complete proteome.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 180 POTENTIAL.
FT SEQUENCE 180 AA; 18610 MW; E0836316B9AD5797 CRC64;

```

```

Query Match 2.0%; Score 7; DB 1; Length 166;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Query Match 2.0%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

SQ SEQUENCE 194 AA; 21960 MW; B38F134AE3DCD0B5 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 194;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 MVAQAMG 314
|||||||
Db 105 MVAQAMG 111

RESULT 10
L0LD_VIBCH STANDARD; PRT; 228 AA.
AC P57066;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Lipoprotein releasing system ATP-binding protein L0LD.
GN L0LD OR VC1883.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=686;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Swinn M.L.,
RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolova M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McComard L., Uterback T., Fleischmann R.D., Nieman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: PART OF AN ATP-DEPENDENT TRANSPORT SYSTEM RESPONSIBLE
FOR THE RELEASE OF LIPOPROTEINS TARGETED TO THE OUTER MEMBRANE
FROM THE INNER MEMBRANE. SUCH A RELEASE IS DEPENDENT OF THE
SORTING-SIGNAL (ABSENCE OF AN ASP AT POSITION 2 OF THE MATURE
LIPOPROTEIN) AND OF L0LA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. L0LD SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AE004263; AAF95031.1; -;
DR TIGR; VC1883; -;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_STP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Complete proteome.
FT NP_BIND 41 48
FT BIND 41 48
SQ SEQUENCE 228 AA; 24690 MW; 315E52CB775E61E6 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 228;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||

Db 119 AKAKAKA 125

RESULT 11
RADCC_BACHD STANDARD; PRT; 232 AA.
AC Q9K8H4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA repair protein radC homolog.
GN RADCC OR BH3032.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RADCC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; AP001517; BAB06751.1; -;
DR InterPro: IPR000445; HHH.
DR InterPro: IPR001405; RadC.
DR Pfam; PF00633; HHH; 1.
DR ProDom; PD007415; RadC; 1.
DR PROSITE; PS01302; RADCC; 1.
KW DNA repair; Complete proteome.
SQ SEQUENCE 232 AA; 26403 MW; 4542153B5C6010D2 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 232;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 152 SINASIV 158
|||||||
Db 151 SINASIV 157

RESULT 12
FIXA_RHILP STANDARD; PRT; 279 AA.
AC Q05559;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-FEB-1995 (Rel. 31, Last annotation update)
GN FIXA protein.
OS Rhizobium leguminosarum (biovar phaseol1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93379046; PubMed=8369342;
RX Michiels J., Vanderleyden J.;

RT "Cloning and sequence of the Rhizobium leguminosarum biovar phaseoli
RT fixa gene.";
RL Blochm. Biophys. Acta 1144:232-233(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN A REDOX PROCESS INVOLVED IN NITROGEN
CC FIXATION.
CC -1- SUBUNIT: FIXA AND FIXB FORM A HETERODIMER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11081; AAA02979.1; -;
DR HSSP: P38117; IEFV.
DR InterPro: IPR000049; ETF_beta.
DR Pfam: PF01012; ETF_beta.1.
DR PRODOM: PD003528; ETF_beta.1.
DR PROSITE: PS01065; ETF_BETA.1.
KW Electron transport; Nitrogen fixation.
SQ SEQUENCE 279 AA; 30083 MW; 6ACB3EABF20D1CE8 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 231 AKAKOI 237
Db 240 AKAKOI 246

RESULT 13
REPL_Bp1 STANDARD; PRT; 281 AA.
AC P19634;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Repl. protein.
GN REPL.
OS Bacteriophage P1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=10678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-69293847; Pubmed-2661831;
RA Hansen E.B.;
RT "Structure and regulation of the lytic replicon of phage P1.";
RL J. Mol. Biol. 207:135-149(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-69293846; Pubmed-2661830;
RA Sternberg N.; Cohen G.;
RT "Genetic analysis of the lytic replicon of bacteriophage P1. II.
RT Organization of replicon elements.";
RL J. Mol. Biol. 207:111-133(1989).
CC -1- FUNCTION: ESSENTIAL FOR L-REPLICON FUNCTION. NECESSARY FOR
CC INITIATION OF DNA REPLICATION FROM THE L-REPLICON (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15639; CAA33662.1; -;
DR EMBL: X15638; CAA33656.1; -;

DR EMBL: X15638; CAA33657.1; ALT_INIT.
DR PIR: S04263; RLBP1.
KW DNA replication.
SQ SEQUENCE 281 AA; 30857 MW; F3C1D033BD756138 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 AKAKA 255
Db 198 AKAKA 204

RESULT 14
OTC_ARCFU STANDARD; PRT; 307 AA.
ID OTC_ARCFU
AC O29013;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine carbamoyltransferase (EC 2.1.3.3) (OTCase).
GN ARCF OR AFI255.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; Pubmed-9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kinkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine -> phosphate
CC + L-citrulline.
CC -1- PATHWAY: ARGININE BIOSYNTHESIS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ARCUSES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001017; AAB89987.1; -;
DR HSSP: O51742; IAI5.
DR TIGR: AF1255;
DR InterPro: IPR002029; Carbamyltransf_asor.
DR Pfam: PF00185; OTCase.1.
DR Pfam: PF02729; OTCase_Nf.1.
DR PRINTS: PR00100; AOTCASE.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; FALSE_NEG.
KW Transferase; Arginine biosynthesis; Complete proteome.
SQ SEQUENCE 307 AA; 34913 MW; DB043BC5A40EB5D CRC64;

Query Match 2.0%; Score 7; DB 1; Length 307;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 ILAALF 266
|||||
DB 166 ILAALF 172

RESULT 15

IN37_SPIOL STANDARD; PRT; 344 AA.
AC P23525;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 37 kDa inner envelope membrane protein, chloroplast precursor (E37).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146572; PubMed=1997321;
RA Drees-Werlingloer U., Fischer K., Wächter E., Link T.A.,
RA Fluegge U.J.;
RT "CDNA sequence and deduced amino acid sequence of the precursor of
RT the 37-kDa inner envelope membrane polypeptide from spinach
RT chloroplasts. Its transit peptide contains an amphiphilic alpha-helix
RT as the only detectable structural element.";
RT Eur. J. Biochem. 195:361-368(1991).
RN [2]
RP SEQUENCE OF 120-141 AND 236-240.
RX MEDLINE=91348205; PubMed=1879527;
RA Block M.A., Joyard J., Douce R.;
RT "Purification and characterization of E37, a major chloroplast
RT envelope protein.";
RL FEBS Lett. 287:167-170(1991).
CC -1- FUNCTION: NOT YET KNOWN. ONE OF THE MAJOR PROTEINS OF THE INNER
CC MEMBRANE OF THE CHLOROPLAST ENVELOPE.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST INNER MEMBRANE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X56963; CAA40283.1; -.
DR PIR: S14409; S14409.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR004033; ubiE_COO5_methyltransf.
DR Pfam: PF01209; ubiE_methyltran; 1.
KW Chloroplast; Transit peptide; Inner membrane.
FT TRANSIT 1 344 CHLOROPLAST.
FT CHAIN 22 344 37 KDA INNER ENVELOPE MEMBRANE PROTEIN.
SQ SEQUENCE 344 AA: 38976 MW: 555E53242B297D7C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 LAKAKAK 254
|||||
DB 154 LAKAKAK 160

0 1 2 3 4 5 6 7 8 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:07 : Search time 20.37 Seconds
(without alignments)
1679.323 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLRAARCMGPFAGLSTG.....RDVQGTDAKDDELDRVAMS 356

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	42.7	357	2 T02246	hypothetical prote
2	14	3.9	334	2 T21562	hypothetical prote
3	13	3.7	515	2 T05863	hypothetical prote
4	11	3.1	311	2 B71689	hypothetical prote
5	11	3.1	312	2 D97756	hypothetical prote
6	11	3.1	354	2 T39599	conserved hypothet
7	9	2.5	160	2 A35148	early protein p12
8	8	2.5	374	2 T18493	hypothetical prote
9	8	2.2	248	2 E87425	transcription regu
10	8	2.2	263	2 G83163	probable amino aci
11	8	2.2	343	2 S09777	hypothetical prote
12	8	2.2	398	2 C96680	hypothetical prote
13	8	2.2	400	2 AG3259	hypothetical prote
14	8	2.2	477	2 T18801	hypothetical prote
15	8	2.2	493	2 D98209	hypothetical prote
16	8	2.2	493	2 AE3077	aldenhyde dehydrat
17	8	2.2	658	2 T20042	aldenhyde dehydrat
18	8	2.2	658	2 T28658	hypothetical prote
19	7	2.0	2124	2 T28658	polyketide synthas
20	7	2.0	47	2 I47015	antigen Mcl (siml
21	7	2.0	95	2 D83325	hypothetical prote
22	7	2.0	95	2 B45519	variant surface g1
23	7	2.0	112	2 F71805	hypothetical prote
24	7	2.0	112	2 D64713	conserved hypothet
25	7	2.0	120	2 A45519	variant surface g1
26	7	2.0	121	2 E72617	hypothetical prote
27	7	2.0	134	1 PWOFE	H+-transporting tw
28	7	2.0	135	2 C72539	hypothetical prote
29	7	2.0	136	2 F95868	conserved hypothet
29	7	2.0	159	2 S65784	ribosomal protein

30	7	2.0	164	2 A84124	DNA-directed RNA p
31	7	2.0	166	1 KRB02B	keratin, 68k type
32	7	2.0	175	2 S43815	hypothetical prote
33	7	2.0	183	2 C97295	stage V sporulatio
34	7	2.0	187	2 H64995	hypothetical prote
35	7	2.0	187	2 B91021	hypothetical prote
36	7	2.0	187	2 A85865	hypothetical prote
37	7	2.0	194	2 F64075	urase accessory p
38	7	2.0	207	2 F83148	hypothetical prote
39	7	2.0	209	2 T17294	hypothetical prote
40	7	2.0	210	2 JC4683	T-cell receptor ga
41	7	2.0	210	2 B75372	hypothetical prote
42	7	2.0	214	2 C82556	hypothetical prote
43	7	2.0	215	2 AE0793	conserved hypothet
44	7	2.0	228	2 A82147	ABC transporter, A
45	7	2.0	232	2 H84028	DNA repair protein

ALIGNMENTS

```
RESULT 1
T02246
hypothetical protein P1.1659_4 - human
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 11-Jan-2000
C:Accession: T02246
R:Jamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.
.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Altix, C.; Andreise, F.; Frank
Submitted to the EMBL Data Library, March 1998
A:Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Ar
A:Description: Sequence analysis of a human P1 clone containing the XRC9 DNA repair
A:Reference number: Z14637
A:Accession: T02246
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <LAM>
A:Cross-references: EMBL:AC004472; NID:q2984582; PIDN:AC07983.1; PID:q2984585
C:Genetics:
A:Map position: 9
A:Insertions: 2/3; 48/3; 82/1; 101/3; 135/3; 194/3; 243/1; 269/3; 312/3
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match      42.7%: Score 152; DB 2; Length 357;
Best Local Similarity 100.0%: Pred. No. 4.9e-142;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 VEAERRRRATVLESEGTRESAINVAEGKKOQIILASEAEKAEOINQAAGESAVALAKAKA 253
      |||||||
Db 195 VEAERRRRATVLESEGTRESAINVAEGKKOQIILASEAEKAEOINQAAGESAVALAKAKA 254

QY 254 KAEIRIRILAAALTOHNGDAASLTVAEOYVSANSKLAKDSNTILPSPNGDVTSMVAQAM 313
      |||||||
Db 255 KAEIRIRILAAALTOHNGDAASLTVAEOYVSANSKLAKDSNTILPSPNGDVTSMVAQAM 314

QY 314 GVGALTKAPVPGPPDLSGSSSRDVGTDAS 345
      |||||||
Db 315 GVGALTKAPVPGPPDLSGSSSRDVGTDAS 346

RESULT 2
T21562
hypothetical protein F30A10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21562
R:Barlow, K.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19442
A:Accession: T21562
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
```

A:Residues: 1-334 <MIL>
 A:Cross-references: EMBL:281072; PIDN:CAB03018.1; GSPDB:GN00019; CESP:F30A10.5
 A:Experimental source: clone F30A10
 C:Genetics:
 A:Gene: CESP.F30A10.5
 A:Map position: 1
 A:Introns: 16/2: 68/3; 124/3; 227/2
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 14; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 FVPOEAMVERMG 53
 ||||||||||||
 Db 47 FVPOEAMVERMG 60

RESULT 3
 T05863

hypothetical protein T29A15.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse ear cress)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 13-Aug-1999
 C:Accession: T05863

R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voelt, M.; Robben, J.; Volckaert, G.; Ho
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: 215455
 A:Accession: T05863

A:Molecule type: DNA
 A:Residues: 1-515 <BEV>
 A:Cross-references: EMBL:AL035602
 A:Experimental source: cultivar Columbia; BAC clone T29A15
 C:Genetics:

A:Map position: 4
 A:Introns: 52/2: 133/3; 183/3; 203/1; 234/1; 267/1; 297/3; 317/3; 394/1; 451/3
 A:Note: T29A15.70

Query Match 3.7%; Score 13; DB 2; Length 515;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTMRSLELG 140
 ||||||||||||
 Db 153 QLAQTMRSLELG 165

RESULT 4
 B71689

hypothetical protein RP328 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: B71689

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, U
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499
 A:Accession: B71689

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-311 <AND>
 A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CA14788.1; PID:g386088
 A:Experimental source: strain Madrid E
 C:Genetics:

A:Gene: RP328
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAQTMRSE 137
 ||||||||||||
 Db 109 TOLAQTMRSE 119

RESULT 5
 D97756

hypothetical protein RC0452 (imported) - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: D97756
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fourminter, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: D97756
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-312 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AL02990.1; PID:g15619524; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC0452
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAQTMRSE 137
 ||||||||||||
 Db 109 TOLAQTMRSE 119

RESULT 6
 T39599

conserved hypothetical protein SPBC16G5.07c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
 C:Accession: T39599
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrall, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21866
 A:Accession: T39599

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-354 <LYN>
 A:Cross-references: EMBL:AL023554; PIDN:CA19027.1; GSPDB:GN00067; SPDB:SPBC16G5.07c
 A:Experimental source: strain 97zh-; cosmid c16G5
 C:Genetics:
 A:Gene: SPDB:SPBC16G5.07c
 A:Map position: 2
 A:Introns: 72/2: 265/2
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 DPKASYGVED 121
 ||||||||||||
 Db 126 DPKASYGVED 136

RESULT 7
 A35148

early protein p12 - phage PRD1
 C:Species: phage PRD1
 C:Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 08-Oct-1999
 C:Accession: A35148; J00186
 R:Gerendasy, D.; Ito, J.
 J. Bacteriol. 172, 1889-1898, 1990

A:Title: Nucleotide sequence and transcription of the right early region of bacteriophage
A:Reference number: A35148; MUID:90202706
A:Accession: A35148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <GER>
A:Cross-references: GB:M30146; NID:g215762; PIDN:AAA2453.1; PID:g215763
R:Palula, T.M.; Savillan, H.; Bamford, D.H.
Gene 85, 53-58, 1989
A:Title: The organization of the right-end early region of bacteriophage PRD1 genome.
A:Reference number: JQ0186; MUID:90152379
A:Accession: JQ0186
A:Molecule type: DNA
A:Residues: 1-160 <PAK>
A:Cross-references: GB:M3428; NID:g215747; PIDN:AAA32447.1; PID:g215748
C:Keywords: DNA binding

Query Match 2.5%; Score 9; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKAA 257
DB 152 AKAKAKAA 160

RESULT 8
T18493
hypothetical protein C0800w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18493
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-374 <LAM>
A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331914; PIDN:CAB11132.1
C:Genetics:
A:Map position: 3
A:Note: C0800w

Query Match 2.5%; Score 9; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 TMRSELGKL 141
DB 174 TMRSELGKL 182

RESULT 9
E87425
transcription regulator, GntR family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 30-Sep-2001
C:Accession: E87425
R:Merz, W.C.; Feldbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolot
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87425
A:Molecule type: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <STO>
A:Cross-references: GB:AE005673; NID:g13422783; PIDN:AAK23401.1; GSPDB:GNC0148
C:Genetics:

A:Gene: CC1420
C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.2%; Score 8; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 256 EAIRILAA 263
DB 65 EAIRILAA 72

RESULT 10
G83163
probable amino acid binding protein PA3865 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83163
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: G83163
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <STO>
A:Cross-references: GB:AE004803; GB:AE004091; NID:g9950035; PIDN:AA607252.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3865
C:Superfamily: lysine-arginine-ornithine-binding protein

Query Match 2.2%; Score 8; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 244 ASAVLAKA 251
DB 151 ASAVLAKA 158

RESULT 11
S09777
hypothetical protein UL14 precursor - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: S09777
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Böhm, R.; Brown, C.M.; Cerny, R.; Horsnell, T
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi
A:Reference number: S09749; MUID:90269039
A:Accession: S09777
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-343 <CHE>
A:Cross-references: EMBL:X17403; NID:g959591; PIDN:CAA35447.1; PID:g959619
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Superfamily: human cytomegalovirus hypothetical protein UL14
C:Keywords: glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-343/Product: hypothetical protein UL14 #status predicted <MAT>
F:281-305/Domain: transmembrane #status predicted <TM>
F:160,255/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.2%; Score 8; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 AAGEASAV 247
 |||||
 DB 323 AAGEASAV 330

RESULT 12

C96680
 hypothetical protein F514.7 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96680
 R:Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Narens, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: C96680
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-398 <STO>
 A:Cross-References: GB:AE005173; NID:g2190544; PIDN:AA60908.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F514.7
 A:Map position: 1

Query Match
 Best Local Similarity 2.2%; Score 8; DB 2; Length 398;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 248 LAKAKAKA 255
 |||||
 DB 77 LAKAKAKA 84

RESULT 13

AG3259
 hypothetical protein BMEI0060 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: AG3259
 R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess,
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3259
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-400 <KUR>
 A:Cross-References: GB:AE008917; PIDN:AAL51242.1; PID:g17981931; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI0060
 A:Map position: 1

Query Match
 Best Local Similarity 2.2%; Score 8; DB 2; Length 400;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 323 PVPGTGRPS 330
 |||||
 DB 21 PVPGTGRPS 28

RESULT 14

T18801
 hypothetical protein ZK131.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T18801; T27746
 R:Percy, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19024
 A:Accession: T18801
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <WIL>
 A:Cross-References: EMBL:Z93373; PIDN:CAB07552.1; GSPDB:GN00020; CESP:ZK131.11
 A:Experimental source: clone C01B9
 R:Steward, C.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z20413
 A:Accession: T27746
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <WIL>
 A:Cross-References: EMBL:Z83245; PIDN:CAB05840.1; GSPDB:GN00020; CESP:ZK131.11
 A:Experimental source: clone ZK131
 C:Genetics:
 A:Gene: CESP:ZK131.11
 A:Map position: 2
 A:Intons: 20/2; 49/2; 113/1; 169/1; 260/2; 375/3; 432/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZK131.11

Query Match
 Best Local Similarity 2.2%; Score 8; DB 2; Length 477;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 KQAOILAS 229
 |||||
 DB 457 KQAOILAS 464

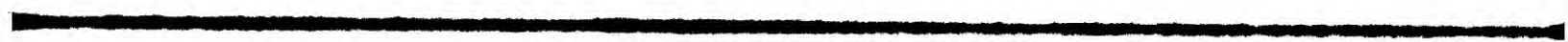
RESULT 15

D98209
 aldehyde dehydrogenase dhds [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: D98209
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Woliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: D98209
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-493 <KUR>
 A:Cross-References: GB:AE007870; PIDN:AAK8198.1; PID:g15159016; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L1241
 A:Map position: linear chromosome
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match
 Best Local Similarity 2.2%; Score 8; DB 2; Length 493;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 AKAFAIRI 260
 |||||
 DB 310 AKAFAIRI 317

Search completed: September 22, 2002, 18:32:59
Job time: 172 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:53 : Search time 39.01 Seconds
(without alignments)
2441.243 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356
Sequence: 1 MLRAARGHWGPFAGELSTG.....RDVGTDAKXDELDRYKMS 356

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 947138 seqs, 267508082 residues

Word size: 0

Total number of hits satisfying chosen parameters: 947138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending_Patents_AA_New*

- 1: /cgn2-6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2-6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2-6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2-6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2-6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2-6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2-6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	354	99.4	356	US-09-898-216-1	Sequence 1, Appl.
2	218	61.2	356	US-09-629-469A-11010	Sequence 11010, A
3	218	61.2	356	US-10-219-793-150	Sequence 150, App
4	202	56.7	306	US-10-219-793-411	Sequence 411, App
5	202	56.7	356	US-60-389-987-1904	Sequence 1904, App
6	152	42.7	357	US-60-389-987-826	Sequence 826, App
7	143	40.2	342	US-60-389-987-768	Sequence 768, App
8	58	16.3	60	US-10-219-793-416	Sequence 416, App
9	58	16.3	60	US-10-219-793-418	Sequence 418, App
10	50	14.0	50	US-10-219-793-420	Sequence 420, App
11	45	12.6	45	US-10-219-793-422	Sequence 422, App
12	45	12.6	45	US-10-219-793-424	Sequence 424, App
13	35	9.8	46	US-10-219-793-426	Sequence 426, App
14	30	8.4	30	US-10-219-793-419	Sequence 419, App
15	28	7.9	28	US-10-219-793-417	Sequence 417, App
16	27	7.6	27	US-10-219-793-413	Sequence 413, App
17	26	7.3	26	US-10-219-793-412	Sequence 412, App
18	26	7.3	26	US-10-219-793-414	Sequence 414, App
19	25	7.0	25	US-10-219-793-421	Sequence 421, App
20	25	7.0	26	US-10-219-793-427	Sequence 427, App
21	24	6.7	24	US-10-219-793-423	Sequence 423, App
22	24	6.7	24	US-10-219-793-425	Sequence 425, App
23	24	6.7	29	US-10-219-793-415	Sequence 415, App
24	17	4.8	29	US-10-219-793-429	Sequence 429, App
25	17	4.8	30	US-10-219-793-428	Sequence 428, App
26	13	3.7	263	US-10-219-999-51693	Sequence 51693, A

27	13	3.7	275	6	US-10-219-999-39902	Sequence 39902, A
28	13	3.7	283	6	US-10-219-999-55341	Sequence 55341, A
29	10	2.8	10	6	US-10-227-616-58	Sequence 58, Appl
30	9	2.5	1109	6	US-10-179-131-6612	Sequence 6612, App
31	8	2.2	15	6	US-10-227-616-59	Sequence 59, Appl
32	8	2.2	43	6	US-10-182-995-20314	Sequence 20314, A
33	8	2.2	43	6	US-10-203-137-26567	Sequence 26567, A
34	8	2.2	43	6	US-10-203-136-26573	Sequence 26573, A
35	8	2.2	43	6	US-10-182-997-25978	Sequence 25978, A
36	8	2.2	43	6	US-10-182-997-18957	Sequence 18957, A
37	8	2.2	43	6	US-10-182-998-10609	Sequence 10609, A
38	8	2.2	43	6	US-10-203-135-25573	Sequence 25573, A
39	8	2.2	43	6	US-10-203-137-26806	Sequence 26806, A
40	8	2.2	43	6	US-10-203-138-10872	Sequence 10872, A
41	8	2.2	43	6	US-10-203-139-25711	Sequence 25711, A
42	8	2.2	187	6	US-10-219-999-48786	Sequence 48786, A
43	8	2.2	255	6	US-10-155-881-10456	Sequence 10456, A
44	8	2.2	471	7	US-60-360-039-11494	Sequence 11494, A
45	8	2.2	471	7	US-60-360-039-14424	Sequence 14424, A

ALIGNMENTS

RESULT 1
US-09-898-216-1
Sequence 1, Application US/09898216

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898,216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-898-216-1

Query Match 99.4%; Score 354; DB 5; Length 356;

Best Local Similarity 100.0%: Pred. No. 0;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLARARGMGPFAGSLSTGFMPRSGRASGGLPRNTVVLFPVQOEAMVVERMGREHRIILE 60
DB 1 MLARARGMGPFAGSLSTGFMPRSGRASGGLPRNTVVLFPVQOEAMVVERMGREHRIILE 60
QY 61 PGLNLLIPVLDRIIRYVQSLKEIIVINPEQSAVTLQIDGVLYLRIMDPYKASGYVE 120
DB 61 PGLNLLIPVLDRIIRYVQSLKEIIVINPEQSAVTLQIDGVLYLRIMDPYKASGYVE 120
QY 121 DPEYAVTQAOITMRSELGKLSXDVFRERESLNASIVDAINQAADCGIRCLREIKRI 180
DB 121 DPEYAVTQAOITMRSELGKLSXDVFRERESLNASIVDAINQAADCGIRCLREIKRI 180
QY 181 HYPVVKESMOMQVEERRKRATVLESECTRESAINVAEGKQAOILASEAKAEQINO 240
DB 181 HYPVVKESMOMQVEERRKRATVLESECTRESAINVAEGKQAOILASEAKAEQINO 240
QY 241 AGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILLPS 300
DB 241 AGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILLPS 300
QY 301 NPGDVTSMVAQAMGYGALTFRAPVGTTPDSSLSSGSRDVGQTDASXDELDRVKMS 356
DB 301 NPGDVTSMVAQAMGYGALTFRAPVGTTPDSSLSSGSRDVGQTDASXDELDRVKMS 356

RESULT 2

US-09-629-469A-11010

; Sequence 11010, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629, 469A

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: JP 1999-300253

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/159,590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/183,322

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 19025

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11010

; LENGTH: 356

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-629-469A-11010

Query Match 61.2%: Score 218; DB 5; Length 356;
Best Local Similarity 99.7%: Pred. No. 8e-197;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 RASSGLPRNTVVLFPVQOEAMVVERMGREHRIILEPGLNLLIPVLDRIIRYVQSLKEIIVN 86
DB 27 RASSGLPRNTVVLFPVQOEAMVVERMGREHRIILEPGLNLLIPVLDRIIRYVQSLKEIIVN 86
QY 87 PQOSAVTLQIDGVLYLRIMDPYKASGYVEDEPEYAVTQAOITMRSELGKLSXKV 146
DB 87 PQOSAVTLQIDGVLYLRIMDPYKASGYVEDEPEYAVTQAOITMRSELGKLSXKV 146
QY 147 FRERESLNASIVDAINQAADCGIRCLREIKRIHYPVVKESMOMQVEERRKRATVLE 206
DB 147 FRERESLNASIVDAINQAADCGIRCLREIKRIHYPVVKESMOMQVEERRKRATVLE 206
QY 207 SECTRESAINVAEGKQAOILASEAKAEQINOAGASAVLAKAKAEAIRILAAAL 266
DB 207 SECTRESAINVAEGKQAOILASEAKAEQINOAGASAVLAKAKAEAIRILAAAL 266
QY 267 QHNGDAAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTFRAPVG 326
DB 267 QHNGDAAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTFRAPVG 326
QY 327 TPDSLSSGSRDVGQTDAS 345
DB 327 TPDSLSSGSRDVGQTDAS 345

RESULT 3

US-10-219-793-150

; Sequence 150, Application US/10219793

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 86 Human Secreted Proteins

; FILE REFERENCE: P2008P1C1

; CURRENT APPLICATION NUMBER: US/10/219, 793

; PRIOR FILING DATE: 2002-08-15

; PRIOR APPLICATION NUMBER: 09/209,462

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: PCT/US98/12125

; PRIOR FILING DATE: 1998-06-11

; PRIOR APPLICATION NUMBER: 60/049,547

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,548

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,549

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,550

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,556

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,606

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,607

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,608

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,609

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,610

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/049,611

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/050,901

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/052,989

; PRIOR FILING DATE: 1997-06-13

; PRIOR APPLICATION NUMBER: 60/051,919

; PRIOR FILING DATE: 1997-07-08

; PRIOR APPLICATION NUMBER: 60/055,984

; PRIOR FILING DATE: 1997-08-18

; PRIOR APPLICATION NUMBER: 60/058,665

; PRIOR FILING DATE: 1997-09-12

; PRIOR APPLICATION NUMBER: 60/058,668

; PRIOR FILING DATE: 1997-09-12

```

; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 150
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-150
```

```

Query Match          61.2%; Score 218; DB 6; Length 356;
Best Local Similarity 99.7%; Pred. No. 8e-197;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 27 RASSGIPRTVYLVFVPOQCAWYVERMGRHRIIEPGLNLIIPVLDIRIVYGSKEIIVNY 86
    |||||||
DB 27 RASSGIPRTVYLVFVPOQCAWYVERMGRHRIIEPGLNLIIPVLDIRIVYGSKEIIVNY 86
QY 87 PEOASVTLNMTLQIDGVLYLRIMDPYKASVGEDEPEYAVTOLAORTMSELGKLSXDKY 146
    |||||||
DB 87 PEOASVTLNMTLQIDGVLYLRIMDPYKASVGEDEPEYAVTOLAORTMSELGKLSXDKY 146
QY 147 FRERESLNASIVDAINQADCMGIRCLRYEIKNDIHVPPRVKESMOMQVEAERKRATVLE 206
    |||||||
DB 147 FRERESLNASIVDAINQADCMGIRCLRYEIKNDIHVPPRVKESMOMQVEAERKRATVLE 206
QY 207 SECTRESAINVAEGKKQAQIILASEAKAEQINQAGEASAVTLAKAKAKAEAIRILAAALT 266
    |||||||
DB 207 SECTRESAINVAEGKKQAQIILASEAKAEQINQAGEASAVTLAKAKAKAEAIRILAAALT 266
QY 267 OHNGDAASLTVAEOYVSAFASKLAKDSNTLLPSPNGDVTSMVAOMGYGALTKAPVPG 326
    |||||||
DB 267 OHNGDAASLTVAEOYVSAFASKLAKDSNTLLPSPNGDVTSMVAOMGYGALTKAPVPG 326
QY 327 TPDSSLSSGSSRDVQGTAS 345
    |||||||
DB 327 TPDSSLSSGSSRDVQGTAS 345
```

```

RESULT 4
US-10-219-793-411
; Sequence 411, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
```

```

; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,556
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-411
```

```

Query Match          56.7%; Score 202; DB 6; Length 306;
Best Local Similarity 100.0%; Pred. No. 8.4e-182;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 144 DKYFRERESLNASIVDAINQADCMGIRCLRYEIKNDIHVPPRVKESMOMQVEAERKRAT 203
    |||||||
DB 94 DKYFRERESLNASIVDAINQADCMGIRCLRYEIKNDIHVPPRVKESMOMQVEAERKRAT 153
```



```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 416
; LENGTH: 60
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-219-793-416

```

```

Query Match          16.3%: Score 58; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.9e-47;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27 RASSGLPRNTVYLFVPOQEAAMVERMGRFRHLEPGINILIPVLDRIRYVQSILKEIYI 84
Db 3 RASSGLPRNTVYLFVPOQEAAMVERMGRFRHLEPGINILIPVLDRIRYVQSILKEIYI 60

RESULT 9
US-10-219-793-418
; Sequence 418, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844

```

```

: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 418
: LENGTH: 60
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-418

Query Match          16.3%; Score 58; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.9e-47;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      85 NVPEGSAYTLDNVTLQIDGVLYLRIMDPYKASYGVDEPEYAVTOLAQTMRSELGKLS 142
Db      1 NVPEGSAYTLDNVTLQIDGVLYLRIMDPYKASYGVDEPEYAVTOLAQTMRSELGKLS 58

RESULT 10
US-10-219-793-420
: Sequence 420, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1C1
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
```

```

: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 420
: LENGTH: 50
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-420

Query Match          14.0%; Score 50; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-39;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      145 KVFRERESINASTIVAINQADCGIRCLRREIKDTHVPRKESKQMOV 194
Db      1 KVFRERESINASTIVAINQADCGIRCLRREIKDTHVPRKESKQMOV 50

RESULT 11
US-10-219-793-422
: Sequence 422, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1C1
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
```

```

: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 422
: LENGTH: 45
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-422
```

```

Query Match          12.6%: Score 45; DB 6; Length 45;
Best Local Similarity 100.0%: Pred. No. 1.4e-34;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 195 EAERRRATVLESEGTRESAINVAEGKQAOILASEAEKAEQINQ 239
Db 1 EAERRRATVLESEGTRESAINVAEGKQAOILASEAEKAEQINQ 45
```

```

RESULT 12
US-10-219-793-424
: Sequence 424, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008PICI
: CURRENT APPLICATION NUMBER: US/10/219,793
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
```

```

: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 424
: LENGTH: 45
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-424
```

```

Query Match          12.6%: Score 45; DB 6; Length 45;
Best Local Similarity 100.0%: Pred. No. 1.4e-34;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 266 TOHNGDAASLTVAEOYVSASFSLAKDSNTILPSPGQVTSNVA 310
Db 1 TOHNGDAASLTVAEOYVSASFSLAKDSNTILPSPGQVTSNVA 45
```

```

RESULT 13
US-10-219-793-426
: Sequence 426, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
```

```
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 426
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-10-219-793-426
```

Query Match 9.8%; Score 35; DB 6; Length 46;

```
Best Local Similarity 100.0%; Pred. No. 3,7e-25;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 QAMGYGALTAKAPVCPDPSLSSGSSRDVGNDAS 345
Db 1 QAMGYGALTAKAPVCPDPSLSSGSSRDVGNDAS 35

RESULT 14
US-10-219-793-419
Sequence 419, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
```



```

; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 419
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-419

Query Match      8.4%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 1,3e-20;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 IDGVLYLRIMDPYKASYGVDPYAVTOLA 130
DB 1 IDGVLYLRIMDPYKASYGVDPYAVTOLA 30

RESULT 15
US-10-219-793-417
; Sequence 417, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,556
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12

```

```

; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 417
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-417

Query Match      7.9%; Score 28; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 9,4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 VPOQEAHVVERMGFRHRLIEPGLNIIIP 68
DB 1 VPOQEAHVVERMGFRHRLIEPGLNIIIP 28

```

Search completed: September 22, 2002, 18:35:49
Job time: 296 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:37 ; Search time 92.04 Seconds
(without alignments)
1361.420 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356
Sequence: 1 MLARAARGHWGPFAGLSTG.....RDVQGTDSXDLELRVYKMS 356

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	218	61.2	356	1	PCT-US98-12125-145
2	218	61.2	356	16	US-09-209-462B-150
3	218	61.2	378	1	PCT-US01-03800A-2351
4	218	61.2	378	1	PCT-US01-04098A-3196
5	218	61.2	3007	1	PCT-US01-04942A-222
6	218	61.2	3019	1	PCT-US01-08656-10594
7	202	56.7	305	16	US-09-298-733-36

8	202	56.7	305	16	US-09-298-733A-36	Sequence 36, App1
9	202	56.7	305	21	US-09-723-594-36	Sequence 36, App1
10	202	56.7	305	21	US-09-724-497-36	Sequence 36, App1
11	202	56.7	306	1	PCT-US98-12125-272	Sequence 272, App
12	202	56.7	306	16	US-09-209-462B-411	Sequence 411, App
13	198	55.6	358	1	PCT-US01-04098A-1228	Sequence 1228, App
14	58	16.3	60	16	US-09-209-462B-416	Sequence 416, App
15	58	16.3	60	16	US-09-209-462B-418	Sequence 418, App
16	51	14.3	51	26	US-60-163-233-2934	Sequence 2934, App
17	50	14.0	50	16	US-09-209-462B-420	Sequence 420, App
18	50	14.0	463	1	PCT-US01-04942A-598	Sequence 598, App
19	45	12.6	45	16	US-09-209-462B-422	Sequence 422, App
20	45	12.6	45	16	US-09-209-462B-424	Sequence 424, App
21	35	9.8	46	16	US-09-209-462B-426	Sequence 426, App
22	33	9.3	56	20	US-09-621-976-4903	Sequence 4903, App
23	33	9.3	56	26	US-60-147-499-4903	Sequence 4903, App
24	30	8.4	30	16	US-09-209-462B-419	Sequence 419, App
25	28	7.9	28	16	US-09-209-462B-417	Sequence 417, App
26	28	7.9	33	26	US-60-163-233-2731	Sequence 2731, App
27	28	7.9	33	26	US-60-164-764-1272	Sequence 1272, App
28	27	7.6	27	1	PCT-US98-12125-274	Sequence 274, App
29	27	7.6	27	16	US-09-209-462B-413	Sequence 413, App
30	26	7.3	26	1	PCT-US98-12125-273	Sequence 273, App
31	26	7.3	26	1	PCT-US98-12125-275	Sequence 275, App
32	26	7.3	26	16	US-09-209-462B-412	Sequence 412, App
33	26	7.3	26	16	US-09-209-462B-414	Sequence 414, App
34	25	7.0	25	16	US-09-209-462B-421	Sequence 421, App
35	25	7.0	26	16	US-09-209-462B-427	Sequence 427, App
36	24	6.7	24	16	US-09-209-462B-423	Sequence 423, App
37	24	6.7	24	16	US-09-209-462B-425	Sequence 425, App
38	24	6.7	70	1	PCT-US98-12125-276	Sequence 276, App
39	24	6.7	70	16	US-09-209-462B-415	Sequence 415, App
40	20	5.6	323	20	US-09-614-150-5415	Sequence 5415, App
41	20	5.6	323	26	US-60-191-637-5433	Sequence 5433, App
42	20	5.6	753	26	US-60-167-217-5529	Sequence 5529, App
43	17	4.8	39	16	US-09-209-462B-429	Sequence 429, App
44	17	4.8	30	16	US-09-209-462B-428	Sequence 428, App
45	13	3.7	265	21	US-09-708-427-57094	Sequence 57094, App

ALIGNMENTS

RESULT 1

PCT-US98-12125-145

Sequence 145, Application PC/TUS9812125

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 86 Human Secreted Proteins

NUMBER OF SEQUENCES: 318

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/12125

FILING DATE: Jan 01, 1900

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36, 373

REFERENCE/DOCKET NUMBER: P2008Complete

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-145

Query Match      61.2%: Score 218; DB 1; Length 356;
Best Local Similarity 99.7%: Pred. No. 4.6e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 RASGSLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Db 27 RASGSLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Oy 87 PEOGAVTLDNVTTLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEOGAVTLDNVTTLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYETKDIHVPKVESMOMOVEAEERRKRAVLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYETKDIHVPKVESMOMOVEAEERRKRAVLE 206
Oy 207 SEGTRESAINVAEKKQKQIILASEEKAEOINOAGESAVALAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEKKQKQIILASEEKAEOINOAGESAVALAKAKAEAIRILAAALT 266
Oy 267 OHNGDAAASLTVAEQVYSAFSLKAKDSNTLLPSNPGDVTSMVAQAMGVYGLTKAPVPG 326
Db 267 OHNGDAAASLTVAEQVYSAFSLKAKDSNTLLPSNPGDVTSMVAQAMGVYGLTKAPVPG 326
Oy 327 TPDSLSSGSSRDVCGTDAS 345
Db 327 TPDSLSSGSSRDVCGTDAS 345

RESULT 2
US-09-209-462B-150
; Sequence 150, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
```

```

; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 150
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-209-462B-150

Query Match      61.2%: Score 218; DB 16; Length 356;
Best Local Similarity 99.7%: Pred. No. 4.6e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 RASGSLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Db 27 RASGSLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Oy 87 PEOGAVTLDNVTTLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEOGAVTLDNVTTLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYETKDIHVPKVESMOMOVEAEERRKRAVLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYETKDIHVPKVESMOMOVEAEERRKRAVLE 206
Oy 207 SEGTRESAINVAEKKQKQIILASEEKAEOINOAGESAVALAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEKKQKQIILASEEKAEOINOAGESAVALAKAKAEAIRILAAALT 266
Oy 267 OHNGDAAASLTVAEQVYSAFSLKAKDSNTLLPSNPGDVTSMVAQAMGVYGLTKAPVPG 326
Db 267 OHNGDAAASLTVAEQVYSAFSLKAKDSNTLLPSNPGDVTSMVAQAMGVYGLTKAPVPG 326
Oy 327 TPDSLSSGSSRDVCGTDAS 345
Db 327 TPDSLSSGSSRDVCGTDAS 345

RESULT 3
```

```
PCT-US01-03800A-2351
; Sequence 2351, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2351
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-03800A-2351
```

```
Query Match          61.2%; Score 218; DB 1; Length 378;
Best Local Similarity 99.7%; Pred. No. 4.9e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 27 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 86
    |||
DB 49 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 108
    |||
QY 87 PEOGSAVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTMRSELGKLSXKV 146
    |||
DB 109 PEOGSAVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTMRSELGKLSXKV 168
    |||
QY 147 FRERESINASTIVAINOAADCGIRCLRYEIKDIHVPKRESOMQVEAERRKRAVLE 206
    |||
DB 169 FRERESINASTIVAINOAADCGIRCLRYEIKDIHVPKRESOMQVEAERRKRAVLE 228
    |||
QY 207 SEGTRESAINVABGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 266
    |||
DB 229 SEGTRESAINVABGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 288
    |||
QY 267 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTAKAPVG 326
    |||
DB 289 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTAKAPVG 348
    |||
QY 327 TPDSLSSGSSRDVOGTGDS 345
    |||
DB 349 TPDSLSSGSSRDVOGTGDS 367
    |||
```

```
RESULT 4
PCT-US01-04098A-3196
; Sequence 3196, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
```

```
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 3196
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-3196
```

```
Query Match          61.2%; Score 218; DB 1; Length 378;
Best Local Similarity 99.7%; Pred. No. 4.9e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 27 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 86
    |||
DB 49 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 108
    |||
QY 87 PEOGSAVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTMRSELGKLSXKV 146
    |||
DB 109 PEOGSAVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTMRSELGKLSXKV 168
    |||
QY 147 FRERESINASTIVAINOAADCGIRCLRYEIKDIHVPKRESOMQVEAERRKRAVLE 206
    |||
DB 169 FRERESINASTIVAINOAADCGIRCLRYEIKDIHVPKRESOMQVEAERRKRAVLE 228
    |||
QY 207 SEGTRESAINVABGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 266
    |||
DB 229 SEGTRESAINVABGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 288
    |||
QY 267 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTAKAPVG 326
    |||
DB 289 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTAKAPVG 348
    |||
QY 327 TPDSLSSGSSRDVOGTGDS 345
    |||
DB 349 TPDSLSSGSSRDVOGTGDS 367
    |||
```

```
RESULT 5
PCT-US01-04942A-222
; Sequence 222, Application PC/TUS0104942A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: PCT/US01/04942A
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 222
; LENGTH: 3007
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04942A-222
```

```
Query Match          61.2%; Score 218; DB 1; Length 3007;
```

Best Local Similarity 99.7%; Pred. No. 3.8e-206;
Matches 318: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 27 RASSGLPRTNTVLEFVPOQEAAMVERMGRFRLIEPGLNITLIPVLDRIIRYQSLKEITVIN 86
Db 27 RASSGLPRTNTVLEFVPOQEAAMVERMGRFRLIEPGLNITLIPVLDRIIRYQSLKEITVIN 86
Oy 87 PEQSAVTLDNVTLQIDGVLTLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEQSAVTLDNVTLQIDGVLTLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Oy 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 QHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAPYVG 326
Db 267 QHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAPYVG 326
Oy 327 TPDSLSGSSSRDVOGTDAS 345
Db 327 TPDSLSGSSSRDVOGTDAS 345
```

RESULT 6
PCT-US01-08656-10594
Sequence 10594, Application PC/TUS0108656

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066
CURRENT APPLICATION NUMBER: PCT/US01/08656
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO: 10594
LENGTH: 3019
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (115)..(144)
OTHER INFORMATION: Band 7 protein family proteins domain identified by EMATRIX.
OTHER INFORMATION: accession number BL01270C, p-value=6.745e-17, raw score of 16.91
NAME/KEY: DOMAIN
LOCATION: (2426)..(2903)
OTHER INFORMATION: ATPases associated with various cellular act domain
OTHER INFORMATION: Identified by Pfam, accession name AAA, E-value=3.3e-182, Pfam
NAME/KEY: misc_feature
LOCATION: (1)..(3019)
OTHER INFORMATION: Xaa - X or * as defined in Table 2

PCT-US01-08656-10594

Query Match 61.2%; Score 218; DB 1; Length 3019;

Best Local Similarity 99.7%; Pred. No. 3.8e-206;
Matches 318: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 27 RASSGLPRTNTVLEFVPOQEAAMVERMGRFRLIEPGLNITLIPVLDRIIRYQSLKEITVIN 86
Db 27 RASSGLPRTNTVLEFVPOQEAAMVERMGRFRLIEPGLNITLIPVLDRIIRYQSLKEITVIN 86
Oy 87 PEQSAVTLDNVTLQIDGVLTLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEQSAVTLDNVTLQIDGVLTLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
```

```
Db 87 PEQSAVTLDNVTLQIDGVLTLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKATYLE 206
Oy 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 QHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAPYVG 326
Db 267 QHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAPYVG 326
Oy 327 TPDSLSGSSSRDVOGTDAS 345
Db 327 TPDSLSGSSSRDVOGTDAS 345
```

RESULT 7
US-09-298-733-36
Sequence 36, Application US/09298733

GENERAL INFORMATION:

APPLICANT: Valenzuela, Dario
APPLICANT: Yuan, Olive
APPLICANT: Hoffman, Heidi
APPLICANT: Hall, Jeff
APPLICANT: Raplejo, Peter
APPLICANT: Adams, R. Mark
APPLICANT: Agostino, Michael J.
APPLICANT: Rechel, Kim
APPLICANT: Howes, Steve H.

TITLE OF INVENTION: Secreted proteins and polynucleotides

FILE REFERENCE: AG199-02PM

CURRENT APPLICATION NUMBER: US/09/298,733

EARLIER FILING DATE: 1999-04-23

EARLIER FILING DATE: 1998-04-24

EARLIER FILING DATE: 1998-05-22

EARLIER FILING DATE: 1998-07-02

EARLIER FILING DATE: 1998-06-11

EARLIER APPLICATION NUMBER: 60/089,163

EARLIER FILING DATE: 1998-06-12

EARLIER APPLICATION NUMBER: 60/091,619

NUMBER OF SEQ ID NOS: 79

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO: 36

LENGTH: 305

TYPE: PRT

ORGANISM: Homo sapiens

US-09-298-733-36

Query Match 56.7%; Score 202; DB 16; Length 305;

Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 202: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 144 DKYFRRESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKAT 203
Db 93 DKYFRRESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERRKAT 152
Oy 204 VLESEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAA 263
Db 153 VLESEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAA 212
Oy 264 ALTOHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAP 323
Db 213 ALTOHNGDAASLTVAEQYVSFAFSKLAQDSNTILLPSNPGDVTSMVAQAMGYGALTKAP 272
Oy 324 VPSTPDSLSGSSSRDVOGTDAS 345
```

Db 273 VPGTPDSLSSGSSRDVQGT DAS 294

RESULT 8
US-09-298-733A-36

; Sequence 36, Application US/09298733A
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapiejko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: 1966.1009-000
; CURRENT FILING DATE: 1999-04-23
; PRIOR FILING DATE: 1998-04-24
; PRIOR FILING DATE: 1998-04-24
; PRIOR FILING DATE: 1998-05-22
; PRIOR FILING DATE: 1998-05-22
; PRIOR FILING DATE: 1998-06-01
; PRIOR FILING DATE: 1998-06-12
; PRIOR FILING DATE: 1998-06-12
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-298-733A-36

Query Match 56.7%; Score 202; DB 16; Length 305;
Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 DKVFRRESLNSIYDAINQADACGICRLRYEIKDIIHPPRVKESMQMVEAERRKRAAT 203
DB 93 DKVFRRESLNSIYDAINQADACGICRLRYEIKDIIHPPRVKESMQMVEAERRKRAAT 152
QY 204 VLESGETRESAINVAEGKKQAOILASEAKAQIINQAGEASAVILAKAKAKAEAIRILAA 263
DB 153 VLESGETRESAINVAEGKKQAOILASEAKAQIINQAGEASAVILAKAKAKAEAIRILAA 212
QY 264 ALTOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAKP 323
DB 213 ALTOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAKP 272
QY 324 VPGTPDSLSSGSSRDVQGT DAS 345
DB 273 VPGTPDSLSSGSSRDVQGT DAS 294

RESULT 9
US-09-723-594-36
; Sequence 36, Application US/09723594
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapiejko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: 1966.1009-002
; CURRENT FILING DATE: US/09/723, 594
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733

; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-723-594-36

Query Match 56.7%; Score 202; DB 21; Length 305;
Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 DKVFRRESLNSIYDAINQADACGICRLRYEIKDIIHPPRVKESMQMVEAERRKRAAT 203
DB 93 DKVFRRESLNSIYDAINQADACGICRLRYEIKDIIHPPRVKESMQMVEAERRKRAAT 152
QY 204 VLESGETRESAINVAEGKKQAOILASEAKAQIINQAGEASAVILAKAKAKAEAIRILAA 263
DB 153 VLESGETRESAINVAEGKKQAOILASEAKAQIINQAGEASAVILAKAKAKAEAIRILAA 212
QY 264 ALTOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAKP 323
DB 213 ALTOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYGALTAKP 272
QY 324 VPGTPDSLSSGSSRDVQGT DAS 345
DB 273 VPGTPDSLSSGSSRDVQGT DAS 294

RESULT 10
US-09-724-497-36
; Sequence 36, Application US/09724497
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapiejko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: 1966.1009-001
; CURRENT FILING DATE: US/09/724,497
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRF
; ORGANISM: Homo sapiens

US-09-724-497-36

Query Match	56.78; Score 202; DB 21; Length 305;
-------------	--------------------------------------

```
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Matches	202;	Conservative	0;	Mismatches	0;	Indels	0;	Caps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

QY	144	DKVFRERESLVAASLYVDALINQAADCMGJRCULRYETKDIHVPRVRESQMOMVEAERRRRAT	203
Db	93	DKVFRERESLVAASLYVDALINQAADCMGJRCULRYETKDIHVPRVRESQMOMVEAERRRRAT	152
QY	204	VLESEBGTRESAINVAEKKKQAOILLASEEKAEBQINQAAGESAVALAKAKAEAIRILAA	263
Db	153	VLESEGTRESAINVAEKKKQAOILLASEEKEQINQAAGESAVALAKAKAEAIRILAA	212
QY	264	ALTOHNGDAASLTVABEYVSAFSEKLAKDSNTILLPSNPBGDVYSMAVQAMGVGALTAP	323
Db	213	ALTOHNGDAASLTVABEYVSAFSEKLAKDSNTILLPSNPBGDVYSMAVQAMGVGALTAP	272
QY	324	VPGRPDSLSSGSSRDVGTGDS	345
Db	273	VPGRPDSLSSGSSRDVGTGDS	294

```

RESULT 11
PCT-US98-12125-272
; Sequence 272, Application PC/TUS9812125
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; NUMBER OF SEQUENCES: 318
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12125
; FILING DATE: Jan 01, 1900
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36, 373
; REFERENCE/DOCKET NUMBER: P2008complete
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 272:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-272

```

Query Match	56.7%	Score 202;	DB 1;	Length 306;
Best Local Similarity	100.0%	Pred. No. 3e-191;		
Matches 202; Conservative	0;	Mismatches	0;	Gaps 0

Oy		144 DKYFRERESLNASTIDAINQADCMWGRCLRYEIKDHPVPRVESMOMOVEAERRKRTAT	203
D6		94 DKYFRERESLNASTIDAINQADCMWGRCLRYEIKDHPVPRVESMOMOVEAERRKRTAT	153

Qy	204	VLESGCTEESKLTINVAEKKQKQOIIILASEKAEKQINQAOAGASAVLAKAKKAKAEIRTLAA	263
Db	154	VLESGCTRESAIIINVAEKKQKQOIIILASEKAEKQINQAOAGASAVLAKAKKAKAEIRTLAA	213
Qy	264	ALTOINGDAALSLVVAEYVAFSKLAKDSNTILLPSNPGDVTSMVAQMGVYCALTKAP	323
Db	214	ALTOINGDAALSLVVAEYVAFSKLAKDSNTILLPSNPGDVTSMVAQMGVYCALTKAP	273
Qy	324	VGTPDLSLSGSSSRDVGCTDAS	345
Db	274	VGTPDLSLSGSSSRDVGCTDAS	295

```

1  US-09-209-462B-411
2  RESULT 2
3  Sequence 411: Application US/09209462B
4  GENERAL INFORMATION:
5  APPLICANT: Rosen et al.
6  TITLE OF INVENTION: 86 Human Secreted Proteins
7  FILE REFERENCE: P2008P1
8  CURRENT APPLICATION NUMBER: US/09/209,462B
9  CURRENT FILING DATE: 1998-12-11
10 PRIOR APPLICATION NUMBER: PCT/US98/12125
11 PRIOR FILING DATE: 1998-06-11
12 PRIOR APPLICATION NUMBER: 60/049,547
13 PRIOR FILING DATE: 1997-06-13
14 PRIOR APPLICATION NUMBER: 60/049,548
15 PRIOR FILING DATE: 1997-06-13
16 PRIOR APPLICATION NUMBER: 60/049,549
17 PRIOR FILING DATE: 1997-06-13
18 PRIOR APPLICATION NUMBER: 60/049,550
19 PRIOR FILING DATE: 1997-06-13
20 PRIOR APPLICATION NUMBER: 60/049,566
21 PRIOR FILING DATE: 1997-06-13
22 PRIOR APPLICATION NUMBER: 60/049,606
23 PRIOR FILING DATE: 1997-06-13
24 PRIOR APPLICATION NUMBER: 60/049,607
25 PRIOR FILING DATE: 1997-06-13
26 PRIOR APPLICATION NUMBER: 60/049,608
27 PRIOR FILING DATE: 1997-06-13
28 PRIOR APPLICATION NUMBER: 60/049,609
29 PRIOR FILING DATE: 1997-06-13
30 PRIOR APPLICATION NUMBER: 60/049,610
31 PRIOR FILING DATE: 1997-06-13
32 PRIOR APPLICATION NUMBER: 60/049,611
33 PRIOR FILING DATE: 1997-06-13
34 PRIOR APPLICATION NUMBER: 60/050,901
35 PRIOR FILING DATE: 1997-06-13
36 PRIOR APPLICATION NUMBER: 60/052,989
37 PRIOR FILING DATE: 1997-06-13
38 PRIOR APPLICATION NUMBER: 60/051,919
39 PRIOR FILING DATE: 1997-07-08
40 PRIOR APPLICATION NUMBER: 60/055,984
41 PRIOR FILING DATE: 1997-08-18
42 PRIOR APPLICATION NUMBER: 60/058,665
43 PRIOR FILING DATE: 1997-09-12
44 PRIOR APPLICATION NUMBER: 60/058,668
45 PRIOR FILING DATE: 1997-09-12
46 PRIOR APPLICATION NUMBER: 60/058,669
47 PRIOR FILING DATE: 1997-09-12
48 PRIOR APPLICATION NUMBER: 60/058,750
49 PRIOR FILING DATE: 1997-09-12
50 PRIOR APPLICATION NUMBER: 60/058,971
51 PRIOR FILING DATE: 1997-09-12
52 PRIOR APPLICATION NUMBER: 60/058,972
53 PRIOR FILING DATE: 1997-09-12
54 PRIOR APPLICATION NUMBER: 60/058,975
55 PRIOR FILING DATE: 1997-09-12
56 PRIOR APPLICATION NUMBER: 60/060,834
57 PRIOR FILING DATE: 1997-10-02
58 PRIOR APPLICATION NUMBER: 60/060,841
59 PRIOR FILING DATE: 1997-10-02
60 PRIOR APPLICATION NUMBER: 60/060,844

```



```

; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-411
```

```
Query Match          56.7%  Score 202: DB 16: Length 306:
Best Local Similarity 100.0%  Pred. No. 3e-191:
Matches 202: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 144 DKVFRRESLNASIVDAINQADCMGIRCLRYEIKDIHVPFRKESMOMQVEARRRRAT 203
DB 94 DKVFRRESLNASIVDAINQADCMGIRCLRYEIKDIHVPFRKESMOMQVEARRRRAT 153
QY 204 VLESEGTRESAINVAEGKKOAOILASEAEKAEQINOAGEASAVLAKAKAKAEAIRILAA 263
DB 154 VLESEGTRESAINVAEGKKOAOILASEAEKAEQINOAGEASAVLAKAKAKAEAIRILAA 213
QY 264 ALTOHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAP 323
DB 214 ALTOHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAP 273
QY 324 VPCTPDSLSSGSSRDYQGTDAAS 345
DB 274 VPCTPDSLSSGSSRDYQGTDAAS 295
```

```
RESULT 13
PCT-US01-04098A-1228
; Sequence 1228, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1228
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1228
```

```
Query Match          55.6%  Score 198: DB 1: Length 358:
Best Local Similarity 99.7%  Pred. No. 3.3e-187:
Matches 298: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 RASSGLPRNTVLFVPEQGEAMVVERKGRFRRLIEPGINILIPVLDRIYQSLKEIYIN 86
DB 27 RASSGLPRNTVLFVPEQGEAMVVERKGRFRRLIEPGINILIPVLDRIYQSLKEIYIN 86
QY 87 PEOSAVTLDNVTLQIDGVLYLRIMDPYKASGYDEPEYATOLAQOTMRELGKLSDKY 146
DB 87 PEOSAVTLDNVTLQIDGVLYLRIMDPYKASGYDEPEYATOLAQOTMRELGKLSDKY 146
QY 147 FRERESLNASIVDAINQADCMGIRCLRYEIKDIHVPFRKESMOMQVEARRRRATVLE 206
DB 147 FRERESLNASIVDAINQADCMGIRCLRYEIKDIHVPFRKESMOMQVEARRRRATVLE 206
QY 207 SEGTRESAINVAEGKKOAOILASEAEKAEQINOAGEASAVLAKAKAKAEAIRILAAALT 266
DB 207 SEGTRESAINVAEGKKOAOILASEAEKAEQINOAGEASAVLAKAKAKAEAIRILAAALT 266
QY 267 QHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAPV 325
DB 267 QHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAPV 325

RESULT 14
US-09-209-462B-416
; Sequence 416, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 1998-06-11
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
```

```

: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060, 834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061, 059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061, 060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 416
: LENGTH: 60
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-209-462B-416

```

```

Query Match      16.3%; Score 58; DB 16; Length 60;
Best Local Similarity 100.0%; Pred. No. 6e-49;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 27 RASSGLPNTVVLVPPQGEAMVVERMGRFHILEPGLNLIPLVDRIYVOSLKEIIV 84
      |||||||
DB 3 RASSGLPNTVVLVPPQGEAMVVERMGRFHILEPGLNLIPLVDRIYVOSLKEIIV 60

```

```

RESULT 15
US-09-209-462B-418
: Sequence 418, Application US/09209462B
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1
: CURRENT APPLICATION NUMBER: US/09/209, 462B
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049, 547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049, 611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050, 901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052, 989

```

```

: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051, 919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055, 984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058, 665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 669
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058, 975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060, 834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060, 865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061, 059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061, 060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 418
: LENGTH: 60
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-209-462B-418

```

```

Query Match      16.3%; Score 58; DB 16; Length 60;
Best Local Similarity 100.0%; Pred. No. 6e-49;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 85 NVPEQSATVLDNVTLQIDGVLYLRIMDPYKASYGVDEPQAVTOLAQTMRSELKLS 142
      |||||||
DB 1 NVPEQSATVLDNVTLQIDGVLYLRIMDPYKASYGVDEPQAVTOLAQTMRSELKLS 58

```

```

Search completed: September 22, 2002, 18:34:49
Job time: 252 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:26:32 ; Search time 15.91 Seconds
(without alignments)
546.544 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLARAARGHMGPFAGELSTG.....RDVQGTDSXDELDPRVKMS 356

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Dackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	99.4	356	1 US-08-781-562-1	Sequence 1, Appl1
2	7	2.0	16	1 US-08-346-849-25	Sequence 25, Appl
3	7	2.0	16	1 US-08-346-849-32	Sequence 32, Appl
4	7	2.0	16	2 US-08-293-284A-25	Sequence 25, Appl
5	7	2.0	16	2 US-08-293-284A-32	Sequence 32, Appl
6	7	2.0	24	4 US-08-993-008A-4	Sequence 4, Appl1
7	7	2.0	29	1 US-08-152-486-3	Sequence 3, Appl1
8	7	2.0	29	1 US-08-303-025-3	Sequence 3, Appl1
9	7	2.0	29	2 US-08-677-304-3	Sequence 3, Appl1
10	7	2.0	35	4 US-08-436-703B-8	Sequence 18, Appl
11	7	2.0	48	4 US-08-995-172-18	Sequence 5, Appl1
12	7	2.0	48	4 US-08-993-008A-5	Sequence 6, Appl1
13	7	2.0	56	4 US-08-993-008A-6	Sequence 2, Appl1
14	7	2.0	60	2 US-08-769-211-2	Sequence 7, Appl1
15	7	2.0	61	4 US-08-995-172-7	Sequence 6, Appl1
16	7	2.0	100	2 US-08-460-890A-64	Sequence 64, Appl
17	7	2.0	100	3 US-08-167-641C-64	Sequence 64, Appl
18	7	2.0	100	4 US-08-460-971A-64	Sequence 64, Appl
19	7	2.0	100	4 US-08-462-040-64	Sequence 13, Appl
20	7	2.0	283	4 US-09-036-987A-13	Sequence 13, Appl
21	7	2.0	283	4 US-09-370-700-13	Sequence 6, Appl1
22	7	2.0	364	4 US-08-891-322-6	Sequence 6, Appl1
23	7	2.0	364	4 US-09-277-019-6	Sequence 6, Appl1
24	7	2.0	381	1 US-08-781-562-6	Sequence 3, Appl1
25	7	2.0	422	2 US-09-067-351-3	Sequence 3, Appl1
26	7	2.0	422	4 US-09-360-490-3	Sequence 3, Appl1
27	7	2.0	422	4 US-09-360-490-3	Sequence 3, Appl1

28	7	2.0	432	3 US-09-075-087-2	Sequence 2, Appl1
29	7	2.0	432	4 US-09-472-971-1	Sequence 1, Appl1
30	7	2.0	433	1 US-08-346-849-2	Sequence 2, Appl1
31	7	2.0	433	2 US-08-293-284A-2	Sequence 2, Appl1
32	7	2.0	553	3 US-09-083-351-2	Sequence 2, Appl1
33	7	2.0	553	4 US-09-083-352-2	Sequence 2, Appl1
34	7	2.0	684	4 US-08-965-762-2	Sequence 2, Appl1
35	7	2.0	693	2 US-08-380-403A-2	Sequence 2, Appl1
36	7	2.0	693	2 US-08-380-403A-5	Sequence 5, Appl1
37	7	2.0	693	2 US-08-895-628-2	Sequence 2, Appl1
38	7	2.0	693	2 US-08-895-628-5	Sequence 5, Appl1
39	7	2.0	713	3 US-09-335-409-11	Sequence 11, Appl
40	7	2.0	713	4 US-09-568-102-11	Sequence 11, Appl
41	7	2.0	713	4 US-09-567-969-11	Sequence 11, Appl
42	7	2.0	713	4 US-09-568-480-11	Sequence 11, Appl
43	7	2.0	713	4 US-09-568-486-11	Sequence 11, Appl
44	7	2.0	713	4 US-09-568-472-11	Sequence 11, Appl
45	7	2.0	740	3 US-08-891-322-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-781-562-1
Sequence 1, Application US/08781562
Patent No. 5763589
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,562
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-08-781-562-1
Query Match 99.4%; Score 354; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 356: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLARAARHGMPFAGLSTGFWPRSGRASSGLPRNTVLEFVPOEAWYVERMGFRHILE 60
 Db 1 MLARAARHGMPFAGLSTGFWPRSGRASSGLPRNTVLEFVPOEAWYVERMGFRHILE 60

Qy 61 PGLNLLIPVLDRIYVOSLKEIVINVPEDSAVTLDNVTLQIDGVLRLIMDPKASVGYE 120
 Db 61 PGLNLLIPVLDRIYVOSLKEIVINVPEDSAVTLDNVTLQIDGVLRLIMDPKASVGYE 120

Qy 121 DPEYAVTOLAQTMRSELGKLSXDKVFRERESLNASIVDAINQADDCMGIRCLREIKNDI 180
 Db 121 DPEYAVTOLAQTMRSELGKLSXDKVFRERESLNASIVDAINQADDCMGIRCLREIKNDI 180

Qy 181 HVPPEKESMQVQVBAERKRKATVLESECTRESAINVAEGKKQAOIILASEAKAQINDA 240
 Db 181 HVPPEKESMQVQVBAERKRKATVLESECTRESAINVAEGKKQAOIILASEAKAQINDA 240

Qy 241 AGEASAVILAKAKAKAARILAAALTOHNGDAASLTVAEOYVSAFSLAKSNTLLPS 300
 Db 241 AGEASAVILAKAKAKAARILAAALTOHNGDAASLTVAEOYVSAFSLAKSNTLLPS 300

Qy 301 NPGDVTSMVAQAMGYGALTKAPVGTPTDLSGSSSRDVOGTDA SXDELDKVKMS 356
 Db 301 NPGDVTSMVAQAMGYGALTKAPVGTPTDLSGSSSRDVOGTDA SXDELDKVKMS 356

RESULT 2
 US-08-346-849-25
 ; Sequence 25, Application US/08346849
 ; Patent No. 5670483
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Shuguang
 ; APPLICANT: Lockshin, Curtis
 ; APPLICANT: Rich, Alexander
 ; APPLICANT: Holmes, Todd
 ; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
 ; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
 ; STREET: Two Millitia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02173-4799
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/346,849
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/973,326
 ; FILING DATE: 28 DECEMBER 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brook, David E.
 ; REGISTRATION NUMBER: 22,592
 ; REFERENCE/DOCKET NUMBER: MIT-6008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 861-6240
 ; TELEFAX: (617) 861-9540
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-346-849-25

Query Match 2.0%; Score 7; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 AKAKAKA 255
 Db 2 AKAKAKA 8

RESULT 3
 US-08-346-849-32
 ; Sequence 32, Application US/08346849
 ; Patent No. 5670483
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Shuguang
 ; APPLICANT: Lockshin, Curtis
 ; APPLICANT: Rich, Alexander
 ; APPLICANT: Holmes, Todd
 ; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
 ; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
 ; STREET: Two Millitia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02173-4799
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/346,849
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/973,326
 ; FILING DATE: 28 DECEMBER 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brook, David E.
 ; REGISTRATION NUMBER: 22,592
 ; REFERENCE/DOCKET NUMBER: MIT-6008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 861-6240
 ; TELEFAX: (617) 861-9540
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-346-849-32

Query Match 2.0%; Score 7; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 AKAKAKA 257
 Db 1 AKAKAKA 7

RESULT 4
 US-08-293-284A-25
 ; Sequence 25, Application US/08293284A
 ; Patent No. 5955343
 ; GENERAL INFORMATION:

```

: APPLICANT: Holmes, Todd
: APPLICANT: Zhang, Shuangang
: APPLICANT: Rich, Alexander
: APPLICANT: DiPersio, C. Michael
: APPLICANT: Lockshin, Curtis
: TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
: TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
: THEREFOR
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/293,284A
: FILING DATE: 22-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/973,326
: FILING DATE: 28-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BROOK, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: MIT-6008A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 25:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-293-284A-25

```

```

Query Match      2.0%; Score 7; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 AKAKAKA 255
|||||
Db 2 AKAKAKA 8

RESULT 5
US-08-293-284A-32
: Sequence 32, Application US/08293284A
: Patent No. 5955343
: GENERAL INFORMATION:
: APPLICANT: Holmes, Todd
: APPLICANT: Zhang, Shuangang
: APPLICANT: Rich, Alexander
: APPLICANT: DiPersio, C. Michael
: APPLICANT: Lockshin, Curtis
: TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
: TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
: THEREFOR
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.

```

```

: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/293,284A
: FILING DATE: 22-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/973,326
: FILING DATE: 28-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BROOK, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: MIT-6008A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-293-284A-32

```

```

Query Match      2.0%; Score 7; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 251 AKAKAKA 257
|||||
Db 1 AKAKAKA 7

```

```

RESULT 6
US-08-993-008A-4
: Sequence 4, Application US/08993008A
: Patent No. 6153596
: GENERAL INFORMATION:
: APPLICANT: Liotta, Dennis C.
: APPLICANT: Petros, John A.
: APPLICANT: Wei, Shioh-Yi
: APPLICANT: Kair, Joan F.
: APPLICANT: Pohl, Jan
: TITLE OF INVENTION: Polycationic Oligomers -
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: Greenlee, Winner and Sullivan
: STREET: 5370 Manhattan Circle, Suite 201
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/993,008A
: FILING DATE: 18-DEC-1997
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/032,436
: FILING DATE: 18-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sullivan, Sally A.
: REGISTRATION NUMBER: 32,064
: REFERENCE/DOCKET NUMBER: 33-95

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-4

Query Match 2.0%; Score 7; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 AKAKAKA 255
|||||
Db 2 AKAKAKA 8

RESULT 7
US-08-152-488-3
Sequence 3, Application US/08152488
Patent No. 5534619
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 512 Springfield Avenue
CITY: Cranford
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07016-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6; ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,488
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: RM-7MG
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-276-3344
TELEFAX: 908-276-5543
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A

TITLE: N/A
PUBLICATION INFORMATION:
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-152-488-3

Query Match 2.0%; Score 7; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 AKAKAKA 255
|||||
Db 4 AKAKAKA 10

RESULT 8
US-08-303-025-3
Sequence 3, Application US/08303025
Patent No. 5614494
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
APPLICANT: Stanley, James C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 150 West Jefferson, Suite 2500
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226-4415
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy diskette 3.5" 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS v.6.22
SOFTWARE: Wordperfect 6.1; ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,025
FILING DATE: 08-SEPT-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06829
FILING DATE: 14-AUG-1992
APPLICATION NUMBER: US 08/152,488
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REFERENCE/DOCKET NUMBER: 7MH-060548-00231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-496-7622
TELEFAX: 313-496-8454
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-303-025-3

Query Match 2.0%; Score 7; DB 1; Length 29;

Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 9
US-08-677-304-3
; Sequence 3, Application US/08677304
; Patent No. 5721212
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,304
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 5721212 Relevant
; TOPOLOGY: No. 5721212 Relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATON INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
US-08-677-304-3

Query Match 2.0%; Score 7; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 10
US-08-436-703B-8
; Sequence 8, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; STREET: Suite 1525
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATON INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
US-08-436-703B-8

Query Match 2.0%; Score 7; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 11
US-08-995-172-18
; Sequence 18, Application US/08995172B
; Patent No. 6218112
; GENERAL INFORMATION:
; APPLICANT: Thatcher, David R

APPLICANT: Wilks, Paula E
TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems
FILE REFERENCE: CAC00026
CURRENT APPLICATION NUMBER: US/08/995,172B
CURRENT FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/033,908
EARLIER FILING DATE: 1996-12-23
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 18
LENGTH: 35
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: UNSURE
LOCATION: (35)
OTHER INFORMATION: Xaa is Cys with Acm sidechain
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-18

Query Match 2.0%; Score 7; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||
DB 7 AKAKAKA 13

RESULT 12
US-08-993-008A-5
Sequence 5, Application US/08993008A
Patent No. 6153596
GENERAL INFORMATION:
APPLICANT: Liotta, Dennis C.
APPLICANT: Petros, John A.
APPLICANT: Wey, Shlow-Jyi
APPLICANT: Karr, Joan F.
APPLICANT: Pohl, Jan
TITLE OF INVENTION: Polycationic Oligomers
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,008A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,436
FILING DATE: 18-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 33-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-5

Query Match 2.0%; Score 7; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||
DB 2 AKAKAKA 8

RESULT 13
US-08-993-008A-6
Sequence 6, Application US/08993008A
Patent No. 6153596
GENERAL INFORMATION:
APPLICANT: Liotta, Dennis C.
APPLICANT: Petros, John A.
APPLICANT: Wey, Shlow-Jyi
APPLICANT: Karr, Joan F.
APPLICANT: Pohl, Jan
TITLE OF INVENTION: Polycationic Oligomers
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,008A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,436
FILING DATE: 18-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 33-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-6

Query Match 2.0%; Score 7; DB 4; Length 56;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 249 AKAKAKA 255
|||||||

Db 2 AKAKAKA 8

RESULT 14

US-08-769-211-2
Sequence 2, Application US/08769211

Patent No. 5830852

GENERAL INFORMATION:

APPLICANT: Thatcher et al.

TITLE OF INVENTION: Improved Pharmaceutical Compositions

TITLE OF INVENTION: Insulin-Receptor Mediated Gene Therapy.

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Inc.

STREET: 75 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1807

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/769,211

FILING DATE: 18-Dec-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Kathleen M. Williams

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3255/02514

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-769-211-2

Query Match 2.0%; Score 7; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255

DB 23 AKAKAKA 29

RESULT 15
US-08-995-172-6
Sequence 6, Application US/08995172B

Patent No. 6218112

GENERAL INFORMATION:

APPLICANT: Thatcher, David R

APPLICANT: Wilks, Paula E

TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems

FILE REFERENCE: CAC00026

CURRENT APPLICATION NUMBER: US/08/995,172B

CURRENT FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: 60/033,908

EARLIER FILING DATE: 1996-12-23

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

LENGTH: 61
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-6

Query Match 2.0%; Score 7; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255

DB 24 AKAKAKA 30

Search completed: September 22, 2002, 18:32:15
Job time: 343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:19:57 ; Search time 104.53 Seconds
(without alignments)
911.058 Million cell updates/sec

Title: US-09-898-216-1
Perfect score: 1767
Sequence: 1 MLARAARGHWGPFRAEGLSTG.....RDVQGTDSXDELDKFKMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 947138 seqs, 267508082 residues

Total number of hits satisfying chosen parameters: 947138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1763	99.8	356	5	US-09-898-216-1
2	1646	93.2	356	5	US-09-629-469A-11010
3	1646	93.2	356	6	US-10-219-793-150
4	1639	92.8	356	7	US-60-389-987-1904
5	1613.5	91.3	342	7	US-60-389-987-768
6	1605	90.8	357	7	US-60-389-987-826
7	1493	84.5	306	6	US-10-219-793-411
8	619.5	35.1	275	6	US-10-219-999-39902
9	578.5	32.7	283	6	US-10-219-999-55341
10	573.5	32.5	263	6	US-10-219-999-51693
11	472	26.7	381	5	US-09-898-216-6
12	466	26.4	305	5	US-09-791-537-98751
13	444.5	25.2	432	5	US-09-602-777A-140
14	376.5	21.3	253	5	US-09-791-537-75409
15	376.5	21.3	576775	5	US-09-458-180-2
16	354.5	20.1	199	5	US-09-791-537-98761
17	354.5	20.1	199	5	US-09-898-216-7
18	295	16.7	60	6	US-10-219-793-418
19	292	16.5	60	6	US-10-219-793-416
20	281.5	15.9	288	5	US-09-898-216-3
21	281.5	15.9	288	6	US-10-223-087-338
22	281.5	15.9	288	6	US-10-223-081-338
23	281.5	15.9	288	6	US-10-223-082-338
24	281.5	15.9	288	6	US-10-223-083-338
25	281.5	15.9	288	6	US-10-223-084-338
26	281.5	15.9	288	6	US-10-223-085-338

27	281.5	15.9	288	6	US-10-223-088-338	Sequence 338, App
28	281.5	15.9	288	6	US-10-223-090-338	Sequence 338, App
29	271	15.3	280	5	US-09-898-216-4	Sequence 4, Appl
30	263.5	14.9	347	1	PCF-US02-13142-3455	Sequence 3455, Ap
31	263.5	14.9	347	6	US-10-128-714-3455	Sequence 3455, Ap
32	261	14.8	50	6	US-10-219-793-420	Sequence 420, App
33	256.5	14.5	347	1	PCF-US02-13142-4455	Sequence 8455, Ap
34	256.5	14.5	347	6	US-10-128-714-8455	Sequence 8455, Ap
35	255.5	14.5	359	6	US-10-179-131-6020	Sequence 6020, Ap
36	231	13.1	70	6	US-10-219-793-415	Sequence 415, App
37	224	12.7	46	6	US-10-219-793-426	Sequence 426, App
38	221	12.5	45	6	US-10-219-793-424	Sequence 424, App
39	211	11.9	45	6	US-10-219-793-422	Sequence 422, App
40	209.5	11.9	383	7	US-60-389-987-853	Sequence 853, App
41	185	10.5	1693	6	US-10-155-881-30739	Sequence 30739, A
42	182.5	10.3	336	6	US-10-219-999-53682	Sequence 53682, A
43	172.5	9.8	329	6	US-10-219-999-57141	Sequence 57141, A
44	171.5	9.7	300	6	US-10-219-999-55568	Sequence 55568, A
45	170.5	9.6	284	6	US-10-219-999-57280	Sequence 57280, A

ALIGNMENTS

RESULT 1
US-09-898-216-1
Sequence 1, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898, 216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-898-216-1
Query Match 99.8%; Score 1763; DB 5; Length 356;

Best Local Similarity 100.0%; Pred. No. 6e-120;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 MLARAARGHMGFPAGLSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 60
Db 1 MLARAARGHMGFPAGLSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 60
Query 61 PGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLRLIMDPYKASYG 120
Db 61 PGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLRLIMDPYKASYG 120
Query 121 DEYAVTOLAQTMRSELGSLSDKVFRRERESLNASTYDAINQADWCGRCLREYEIK 180
Db 121 DEYAVTOLAQTMRSELGSLSDKVFRRERESLNASTYDAINQADWCGRCLREYEIK 180
Query 181 HVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKQAQIILASEAKAEQIN 240
Db 181 HVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKQAQIILASEAKAEQIN 240
Query 241 AGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 300
Db 241 AGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 300
Query 301 NPGDVTSMVAQMGVYALTAPVPGTPTDLSGSSRDVQGTDA SXDELDRYKMS 356
Db 301 NPGDVTSMVAQMGVYALTAPVPGTPTDLSGSSRDVQGTDA SXDELDRYKMS 356

RESULT 2
US-09-629-469A-11010
; Sequence 11010, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; FILE REFERENCE: ISOGAI, TAKAO
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11010
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-11010

Query Match 93.2%; Score 1646; DB 5; Length 356;
Best Local Similarity 95.3%; Pred. No. 1.9e-111;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

Query 1 MLARAARGHMGFPAGLSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
Db 1 MLARAARGHMGFPAGLSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
Query 59 LEPGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLRLIMDPYKASYG 118
Db 59 LEPGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLRLIMDPYKASYG 118
Query 119 VEDPEYAVTOLAQTMRSELGSLSDKVFRRERESLNASTYDAINQADWCGRCLREYEIK 178
Db 119 VEDPEYAVTOLAQTMRSELGSLSDKVFRRERESLNASTYDAINQADWCGRCLREYEIK 178
Query 179 DIHVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKQAQIILASEAKAEQIN 238
Db 179 DIHVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKQAQIILASEAKAEQIN 238
Query 239 QAAGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 298
Db 239 QAAGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 298
Query 299 PSNPGDVTSMVAQMGVYALTAPVPGTPTDLSGSSRDVQGTDA SXDELDRYKMS 356
Db 299 PSNPGDVTSMVAQMGVYALTAPVPGTPTDLSGSSRDVQGTDA SXDELDRYKMS 356

RESULT 3
US-10-219-793-150
; Sequence 150, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; FILE REFERENCE: P2008PICI
; CURRENT FILING DATE: US/10/219,793
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12


```

Oy 77 QSLKEIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 136
Db 63 QSLKEIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 122
Oy 137 ELGLSLAKVYRERESLNAVDAINQADMGIRCLREYKIDHPPRVESQMOMEYA 196
Db 123 ELGLSLDKVFRERESLNAVDAINQADMGIRCLREYKIDHPPRVESQMOMEYA 182
Oy 197 ERRKATVLESEGTRESAINVAEGKKQAOILASEAKAQINOAGSAVLAQAKAKAE 256
Db 183 ERRKATVLESEGTRESAINVAEGKKQAOILASEAKAQINOAGSAVLAQAKAKAE 242
Oy 257 AIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNGDYTSVVAQAMGY 316
Db 243 AIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNGDYTSVVAQAMGY 302
Oy 317 GALTAPVPGTPDLSGSSRDVQGTDA SXDELDVRKMS 356
Db 303 GALTAPVPGTPDLSGSSRDVQGTDA SXDELDVRKMS 342

RESULT 6
US-60-389-987-826
Sequence 826, Application US/60389987
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 650088,465P2
CURRENT APPLICATION NUMBER: US/60/389,987
CURRENT FILING DATE: 2002-06-17
NUMBER OF SEQ ID NOS: 3025
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 826
LENGTH: 357
TYPE: PRT
ORGANISM: Homo sapiens
US-60-389-987-826

Query Match          90.8%: Score 1605; DB 7; Length 357;
Best Local Similarity 93.8%: Pred No. 1,8e-108;
Matches 332; Conservative 2; Mismatches 4; Indels 16; Gaps 2;

Oy 17 LSTGFWRPSGRASSGLPRTVYLVFPOQEAVMYERMGFRHILPEJINILIPVLDRIYV 76
Db 6 LASGRAPR--RASSGLPRTVYLVFPOQEAVMYERMGFRHILPEJINILIPVLDRIYV 63
Oy 77 QSLKEIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 136
Db 64 QSLKEIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 123
Oy 137 ELGLSLDKVFRERESLNAVDAINQADMGIRCLREYKIDHPPRVESQMOMEYA 196
Db 124 ELGLSLDKVFRERESLNAVDAINQADMGIRCLREYKIDHPPRVESQMOMEYA 183
Oy 194 -----VEARRKATVLESEGTRESAINVAEGKKQAOILASEAKAQINOAG 242
Db 184 KEGMEKGLRAPVEARRKATVLESEGTRESAINVAEGKKQAOILASEAKAQINOAG 243
Oy 243 EASAVLAKAKAKAERITLAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNP 302
Db 244 EASAVLAKAKAKAERITLAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNP 303
Oy 303 GDVTSVVAQAMGYGALTAPVPGTPDLSGSSRDVQGTDA SXDELDVRKMS 356
Db 303 GDVTSVVAQAMGYGALTAPVPGTPDLSGSSRDVQGTDA SXDELDVRKMS 356
```

```

Db 304 GDVTSVVAQAMGYGALTAPVPGTPDLSGSSRDVQGTDA SXDELDVRKMS 357

RESULT 7
US-10-219-793-411
Sequence 411, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
```



```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-219-793-411

```

```

Query Match          84.5%; Score 1493; DB 6; Length 306;
Best Local Similarity 99.3%; Pred. No. 1.9e-100;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 51 RMGRFRRIIEPGIINILIPYLDRIYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIM 110
DB 1 RMGRFRRIIEPGIINILIPYLDRIYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIM 60
QY 111 DPKASYGVEDEPYAVTOLAQTMRSELGKLSKDXYFRRESINASTIVDAINQADCWGI 170
DB 61 DPKASYGVEDEPYAVTOLAQTMRSELGKLSKDXYFRRESINASTIVDAINQADCWGI 120
QY 171 RCLRYEIKIIVHPRVKESMOMOVEARRRRATVLESEGTRESAINVAEGKQOILLASE 230
DB 121 RCLRYEIKIIVHPRVKESMOMOVEARRRRATVLESEGTRESAINVAEGKQOILLASE 180
QY 231 AEKAEQINQAGEASAVLAKAKAKAEIRILAAALTQHNQDAASLTVAEQYSAFSKLA 290
DB 181 AEKAEQINQAGEASAVLAKAKAKAEIRILAAALTQHNQDAASLTVAEQYSAFSKLA 240
QY 291 KQSNITLFSNPGDVTSMVAQAMGYGALTAPVPGTPOSLSGSSRDVQGTDAKXDEEL 350
DB 241 KQSNITLFSNPGDVTSMVAQAMGYGALTAPVPGTPOSLSGSSRDVQGTDAKXDEEL 300
QY 351 DRYKMS 356
DB 301 DRYKMS 306

```

```

RESULT 8
US-10-219-999-39902

```

```

; Sequence 39902, Application US/10219999

```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 39902
; LENGTH: 275
; TYPE: prt
; ORGANISM: Zea mays
US-10-219-999-39902

```

```

Query Match          35.1%; Score 619.5; DB 6; Length 275;
Best Local Similarity 48.2%; Pred. No. 4.1e-37;
Matches 131; Conservative 47; Mismatches 77; Indels 17; Gaps 2;

```

```

QY 100 QIDGVLYLRIMDPYKASYGVEDEPYAVTOLAQTMRSELGKLSKDXYFRRESINASTIV 159
DB 1 QIDGVLYLRIMDPYKASYGVEDEPYAVTOLAQTMRSELGKLSKDXYFRRESINASTIV 60
QY 160 AINQADCWGIRCLRYEIKIIVHPRVKESMOMOVEARRRRATVLESEGTRESAINVAE 219
DB 1 AINQADCWGIRCLRYEIKIIVHPRVKESMOMOVEARRRRATVLESEGTRESAINVAE 219

```

```

DB 61 AINEATDWCIGKICIRYEIRIDINPPACIRQAMENQAEERKKRAQILLESQKQOILLSE 120
QY 220 GKQKQOILLSEAKAEQINQAGEASAVLAKAKAKAEIRILAAALTQHNQDAASLTVA 279
DB 121 GKRTAQILLESQKQOILLSEAKAEQINQAGEASAVLAKAKAKAEIRILAAALTQHNQDAASLTVA 180
QY 280 EQYVSAFSKLAQSNITLFSNPGDVTSMVAQAMGYGALTAPVPGTPOSLSGSSRDVQGTDAKXDEELDRVK 354
DB 181 EQYVSAFSKLAQSNITLFSNPGDVTSMVAQAMGYGALTAPVPGTPOSLSGSSRDVQGTDAKXDEELDRVK 240
QY 324 -VPGTPOSLSGSSRDVQGTDAKXDEELDRVK 354
DB 241 ETPAPSSSEASKTPPLEEADSNQTFSLQRPK 272

```

```

RESULT 9
US-10-219-999-55341

```

```

; Sequence 55341, Application US/10219999

```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 55341
; LENGTH: 283
; TYPE: prt
; ORGANISM: Zea mays
US-10-219-999-55341

```

```

Query Match          32.7%; Score 578.5; DB 6; Length 283;
Best Local Similarity 53.4%; Pred. No. 4e-34;
Matches 119; Conservative 37; Mismatches 50; Indels 17; Gaps 2;

```

```

QY 14 AEGLSGTFMFRSGRASSGL-----PRNTYVLYFPQOEAENVBERMGRFHILRGLNILL 67
DB 72 APALSRFSFERNRDSMSMEDPEPPVNMGSIVPEKKAVVERFGRKYLTSQGFHLI 131
QY 68 PVLDRIRYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIMDPYKASYGVEDEPYAVT 127
DB 132 PAVDRIRYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIMDPYKASYGVEDEPYAVT 191
QY 128 QLAQTMRSELGKLSKDXYFRRESINASTIVDAINQADCWGIRCLRYEIKIIVHPRVK 187
DB 192 QLAQTMRSELGKLSKDXYFRRESINASTIVDAINQADCWGIRCLRYEIKIIVHPRVK 251
QY 188 ESMQOMVEARRRRATVLESEGTRESAINVAEGKQOILLASE 230
DB 252 QAMEQOAEARRRRATVLESEGTRESAINVAEGKQOILLASE 283

```

```

RESULT 10
US-10-219-999-51693

```

```

; Sequence 51693, Application US/10219999

```

```

; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

```


OY 90 SAVTLDNVTLOIGVLYLRIMDPYKASYGVEDPEYAVTOLAQTMRSELGKLSXDKVERE 149
: | | | | : | : | | | | | | | | | | | | | | | | : | : | :
DB 74 DVITKDNVTVOVDAYVFRVDPVKAIVEDEYFATSOIAQTTLRSVCGEALDELDSQ 133
: | | | | : | : | | | | | | | | | | | | | | | | : | : | :
OY 150 RESLNASIVDAINQAADCIGIRCLREYKDIHVPPRVKESMOMQVBAERRRKRAVLESEG 209
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 134 REKINMKLOEIIDROTDPGCVKIVAVELKKIDLPEELRKALARQAERERERRAKTIISAA 193
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY 210 TRESAINVAEGKKOAOILASE 230
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 194 EYQAAOKLEA---ARILAOE 211
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 15
US-09-458-180-2
: Sequence 2, Application US/09458180
: GENERAL INFORMATION:
: APPLICANT: DIVERSA CORPORATION
: APPLICANT: Swanson, Ronald V.
: APPLICANT: Short, Jay M.
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF THE AQUIFEX
: TITLE OF INVENTION: AEOLICUS GENOME, FRAGMENTS THEREOF, AND USES THEREOF
: FILE REFERENCE: DIVER1330
: CURRENT APPLICATION NUMBER: US/09/458,180
: CURRENT FILING DATE: 1999-12-08
: PRIOR APPLICATION NUMBER: US 08/895,611
: PRIOR FILING DATE: 1997-07-16
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 576775
: TYPE: PRF
: ORGANISM: Aquifex aeolicus
US-09-458-180-2

Query Match 21.3%; Score 376.5; DB 5; Length 576775;
Best Local Similarity 39.3%; Pred. No. 4e-15;
Matches 79; Conservative 47; Mismatches 64; Indels 11; Gaps 3;

OY 37 VVLF-----VPOQEAHVVERGRHRLERLEGLNLIPVLDRIYVOSLKEIVINVPEQ 89
: | | | | : | : | | | | | | | | | | | | | | | : | : | :
DB 553422 VVLFPLASAIKVIPEYERAVFRIGAKGPGILIVIPIDIRIVR-SLRTVTLDPVTO 553480
: | | | | : | : | | | | | | | | | | | | | | | : | : | :
OY 90 SAVTLDNVTLOIGVLYLRIMDPYKASYGVEDPEYAVTOLAQTMRSELGKLSXDKVERE 149
: | | | | : | : | | | | | | | | | | | | | | | : | : | :
DB 553481 DVITKDNVTVOVDAYVFRVDPVKAIVEDEYFATSOIAQTTLRSVCGEALDELDSQ 553540
: | | | | : | : | | | | | | | | | | | | | | | : | : | :
OY 150 RESLNASIVDAINQAADCIGIRCLREYKDIHVPPRVKESMOMQVBAERRRKRAVLESEG 209
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 553541 REKINMKLOEIIDROTDPGCVKIVAVELKKIDLPEELRKALARQAERERERRAKTIISAA 553600
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY 210 TRESAINVAEGKKOAOILASE 230
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 553601 EYQAAOKLEA---ARILAOE 553618
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

Search completed: September 22, 2002, 18:30:32
Job time: 635 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:26:03 ; Search time 44.59 Seconds

(without alignments)
767.163 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767

Sequence: 1 MLRAARGHWGPFPAEGLSTG.....RDVQGTDSXDELDPRVYKMS 356

Scoring table: BLOSUM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1605	90.8	357	2 T02246	hypothetical prote
2	957.5	54.2	334	2 T21562	hypothetical prote
3	861.5	48.8	354	2 T39599	conserved hypotet
4	826	46.7	515	2 T05863	hypothetical prote
5	811.5	45.9	312	2 D97756	hypothetical prote
6	806	45.6	311	2 B71689	hypothetical prote
7	722	40.9	374	2 T18493	hypothetical prote
8	716	40.5	315	2 H81907	probable periplasm
9	712.5	40.3	315	2 F81107	stomatoln/Mec-2 fam
10	499.5	28.3	321	2 S74505	erthyrocyte band 7
11	497	28.1	344	2 D75311	conserved hypotet
12	496.5	28.1	337	2 A82838	conserved hypotet
13	486	27.5	310	2 E87667	band 7/Mec-2 faml
14	482.5	27.3	322	2 A12114	hypothetical prote
15	477	27.0	304	2 A80375	conserved hypotet
16	475	26.9	319	2 T36857	conserved hypotet
17	472	26.7	381	2 C70711	hypothetical prote
18	468	26.5	305	2 A80564	probable membrane
19	467.5	26.5	306	2 F82256	conserved hypotet
20	466	26.4	305	2 H64779	probable membrane
21	466	26.4	305	2 H90697	probable proteins
22	466	26.4	305	2 B85548	probable proteins
23	465	26.3	350	2 A98264	hypothetical 33.7K
24	465	26.3	350	2 A83020	conserved hypotet
25	460	26.0	298	2 A71159	probable membrane
26	460	26.0	299	2 F75003	stomatoln-like prot
27	457	25.9	318	2 C69192	stomatoln-like prot
28	454.5	25.7	328	2 A13511	stomatoln like prot
29	444.5	25.2	374	2 C87134	conserved hypotet

30	442.5	25.0	322	2 A97030	membrane protease
31	423.5	24.0	305	2 E72322	conserved hypotet
32	417.5	23.6	392	2 G84191	hypothetical prote
33	412	23.3	309	2 D90520	conserved hypotet
34	376.5	21.3	253	2 F70378	erythrocyte band 7
35	369	20.9	278	2 AC2354	hypothetical prote
36	362	20.5	249	2 C69427	membrane protein h
37	354.5	20.1	199	2 C64403	membrane protein h
38	345.5	19.6	278	2 D72522	probable erythrocy
39	339	19.2	297	2 G69104	stomatoln-like prot
40	327.5	18.5	266	2 C71027	probable erythrocy
41	324.5	18.4	264	2 A83590	probable stomatin-
42	324	18.3	256	2 B95981	probable stomatin-
43	323.5	18.3	262	2 T34691	hypothetical prote
44	321.5	18.2	268	2 B75107	stomatoln-like prot
45	310	17.5	267	2 H90388	erythrocyte band 7

ALIGNMENTS

RESULT 1

T02246
hypothetical protein P1.11659_4 - human

C:Species: Homo sapiens (man)

C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 11-Jan-2000

C:Accession: T02246

R:lamerdin, J.E.; McGready, P.M.; Skowronski, E.; Adamson, A.W.; Burthart-Schultz, K.

.. Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Atlix, C.; Andreise, T.; Trank

submitted to the EMBL Data Library, March 1998

A:Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Krommiller, B.; Ar

A:Description: Sequence analysis of a human P1 clone containing the XRC9 DNA repair

A:Reference number: 214637

A:Accession: T02246

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-357 <L>AM>

A:Cross-References: EMBL:NC004472; NID:g2984582; PIDN:AAC07983.1; PID:g2984585

A:Map position: 9

A:introns: 2/3; 48/3; 82/1; 101/3; 135/3; 194/3; 243/1; 269/3; 312/3

A:Superfamily: erythrocyte band 7 integral membrane protein

Query Match	Score	Length	DB	2:	Indels	Gaps
Best Local Similarity	93.8%	Pred. No. 2.9e-98;				
Matches	332;	Conservative	2;	Mismatches	4;	Indels 16; Gaps 2;
QY	17	LSTGFWRSGRASGLPRTVTLFVPOQAWYERMGREFRILEPGLNILLIPVLRIRYV	76			
DB	6	LASGRAPR--RASGLPRTVTLFVPOQAWYERMGREFRILEPGLNILLIPVLRIRYV	63			
QY	77	OSLKEIVNPEQSAVTLQIDGVLYLRIMDPYKASYGVDEPVAVTOLAOTMRS	136			
DB	64	OSLKEIVNPEQSAVTLQIDGVLYLRIMDPYKASYGVDEPVAVTOLAOTMRS	123			
QY	137	ELGKLSXDKVFERESLNSIVDAINQADCGIRCLREIKOIHVPYKESMOMQ---	193			
DB	124	ELGKLSXDKVFERESLNSIVDAINQADCGIRCLREIKOIHVPYKESMOMQVCA	183			
QY	194	-----YEAERRKRAVLESEGTRESAINVAEGKKQAQIILASEAKAEQIQNAG	242			
DB	184	KEGWKGLAPYEAERKRAVLESEGTRESAINVAEGKKQAQIILASEAKAEQIQNAG	243			
QY	243	ESAVYAKAKAKAEATRIITAAALTOHNGDAASLYIAEOTVSAFSLADSNITLLPSNP	302			
DB	244	ESAVYAKAKAKAEATRIITAAALTOHNGDAASLYIAEOTVSAFSLADSNITLLPSNP	303			
QY	303	GDVTSVAQAMGVYGLTKAPVPGTDSLSGSSSRVVOGTDSXDELDPRVYKMS	356			
DB	304	GDVTSVAQAMGVYGLTKAPVPGTDSLSGSSSRVVOGTDSXDELDPRVYKMS	357			

A:Accession: D75311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-344 <NH2>
A:Cross-references: GB:AE002048; GB:AE000513; NID:g6459929; PIDN:AAF11687.1; PID:g645993
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2143
A:Map position: 1
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 28.1%; Score 497; DB 2; Length 344;
Best Local Similarity 37.2%; Pred. No. 1.9e-25;
Matches 108; Conservative 67; Mismatches 103; Indels 12; Gaps 5;
QY 41 VQOZAMVYERGRFRHLEPGLNLLPVLDRIRYVQSLKEIYINPEQSAVTLDNVLT 100
DB 25 VQOZAMVYERGRFRHLEPGLNLLPVLDRIRYVQSLKEIYINPEQSAVTLDNVLT 84
QY 101 IDGVYLRIMDPYKASVGEDEPEYAVTOLAOTMRSELGSLKDXKVERRESINASTIVA 160
DB 85 VDAVYFYOVLDAKASVYERGRFRHLEPGLNLLPVLDRIRYVQSLKEIYINPEQSAVTLDNVLT 144
QY 161 INQADCMGIRCLRYEIKDHPVPRVYKESQMOQVEAEERKRAVLESECTFESAINVAG 220
DB 145 VDEAREPMGVKVTREIVKDKIPKPADLVASMAROMKAREKRNILIDAEGFQAILLKADG 204
QY 221 KQAOILASEAKAEQINQAA-GEASAVLAKAKAKAEARITLAAALTOHNGDAASLTVA 279
DB 205 EKQAVANLKAEGK-----QASFMSEARERRRAQAEAEARVYVSOALAGCNVQ-AYNYFLA 258
QY 280 EGYVSFSLKADSN--TILPSNPGDVTSMVAQAMGVGALTKAPVPGT 327
DB 259 QGYVEALRVASAPNOKTILPI---EATSIIGSLGIGLAEVAKETPGS 305

RESULT 12
A82838
conserved hypothetical protein XF0185 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: A82838
R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82838
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <SIM>
A:Cross-references: GB:AE003872; GB:AE003849; NID:g9104975; PIDN:AAF2998.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A-
Biones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0185
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 28.1%; Score 496.5; DB 2; Length 337;
Best Local Similarity 36.2%; Pred. No. 1.9e-25;
Matches 119; Conservative 67; Mismatches 116; Indels 27; Gaps 8;
QY 30 SGLPRNTV-----VLFVPOQZAMVYERGRFRHLEPGLNLLPVLDRIR 74
DB 19 SMLPNVLAIVLVAGVILLESFVIMVPOGYEWTVEKFGRYTDMKPLGFILPLISVG 78
QY 75 YVQSLKEIYINPEQSAVTLDNVLTIDGVYLRIMDPYKASVGEDEPEYAVTOLAOTM 134
DB 79 RYVSMEOVLAVPSQEVITKDNAGVGVDFEYVLDKAKAAEVANLEIMATLVITNI 138
QY 135 RSELGSLKDXKVERRESINASTIVAIDANOACMGIRCLRYEIKDHPVPRVYKESQMOQV 194
DB 139 RYVGSIDPDESQRETTINAKLSVEHATSPWGKVTIRIDKIDPPHNLAESMOQK 198
QY 195 EAEERRRATVLESECTFESAINVAGKQAOILASEAKAEQINQAA-GEASAVLAKAKAKAE 254
DB 199 MEQTRRATILAEAGVROSAILRADGEKQAAVYEAERKEAFAFRDA--EARETLAEAEAK 256
QY 255 AEARITLAAALTOHNGDAASLTVAEOYVSAPSKLADSN--TILPSNPGDVTSMVAQAM 312
DB 257 --ATRLISEAISQGNVQ-AVNIFVAKYVAERKEIATAPNOKFILPMETSGISIA-- 311
QY 313 MGYYGALTKAPVPGTSPDLSGSSRPDVG 341
DB 312 -GI-ADLTREAL-GKSDHATPSQREHRG 337

RESULT 13
E87667
band 7/mec-2 family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: E87667
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87667
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <STO>
A:Cross-references: GB:AE005673; NID:g13425081; PIDN:AAK25337.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3375
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 27.5%; Score 486; DB 2; Length 310;
Best Local Similarity 35.5%; Pred. No. 8.5e-25;
Matches 107; Conservative 79; Mismatches 97; Indels 18; Gaps 7;
QY 36 TVVLEFVPOQZAMVYERGRFRHLEPGLNLLPVLDRIRYVQSLKEIYINPEQSAVTL 95
DB 5 SAKIIVPQGRREFVYERGRFRHLEPGLNLLPVLDRIRYVQSLKEIYINPEQSAVTL 64
QY 96 NVTLOIDGVYLRIMDPYKASVGEDEPEYAVTOLAOTMRSELGSLKDXKVERRESIN 155
DB 65 NNSVKDAIVLFIOVMDAAAVRVDMIMYATLOAOTNLTFTVGAEMLELVLSQRAINS 124
QY 156 STYDANQADCMGIRCLRYEIKDHPVPRVYKESQMOQVEAEERKRAVLESECTFESAI 215
DB 125 RLSTIDHTATGPGVAVVATRIKIDLPVPADITNAMRQKAEERAVTEAGGESQAOI 184
QY 216 NVAEGKQAOILASEAKAEQINQAA-GEASAVLAKAKAKAEARITLAAALTOHNGDA-AA 274
DB 185 ARAEGKQAOILASEAKAEQINQAA-GEASAVLAKAKAKAEARITLAAALTOHNGDA-AA 274
QY 275 SLTVAEOYVSAPSKLAK--DSNTILPSNPGDVTSMVAQAMGVGALTK-----ADVP 327

QY	156	SYVDIAINQADQWGIRCLBYELKIHVPRVYESQMOMQVEARRRRRATVLEEGTRESAI	215
Db	138	LRLRELDISTDPGWKAVTVRELRLDIPPSKAVLSDSELMQTAERKKRAALITTEGGORDSAI	197
QY	216	NVAEGKQAOIILASAEKAEQINOAGAENSAVALAKAKAEKIRILIAALTOHN-GDAA	274
Db	198	NSAQDGAARVLEAEAKKKAAITLMAEAGQKKVLEAKKATVAEALSTLTERKSSDNHAREL	257
QY	275	SITVAEOYVSAFSKLAKDSNTILPESNEDDTSMVAQAMGVG	317
Db	258	OFLAADOYLMGTITGSSDSKVMFLPRLNITLSTEGVRSIVG	300
RESULT 2			
YE88_MYCTU	STANDARD: PRT: 381 AA.		
AC	P71768:		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical 41.2 kDa protein RV1488.		
GN	RV1488 OR MT1533.2 OR MGC277.09.		
OS	Mycobacterium tuberculosis.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxID=1773;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=H37RV;		
RX	MEDLINE=98295987; PubMed=9634230;		
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
RA	Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,		
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,		
RA	Davies R., Davlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,		
RA	Hornsbly T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,		
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,		
RA	Rutler S., Seeger K., Skelton S., Squares S., Squares R.,		
RA	Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;		
RT	"Deciphering the biology of Mycobacterium tuberculosis from the		
RT	complete genome sequence."		
RL	Nature 393:537-544(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CDC 1551 / Oshkosh;		
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,		
RA	Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,		
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,		
RA	Delcher A., Utterback T., Weisman J., Kouri H., Gill J., Mikula A.,		
RA	Bisbal W.;		
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and		
RT	laboratory strains."		
RL	Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: Z79701; CABO2038.1; -		
DR	EMBL: AE007022; AAK45800.1; -		
DR	TIGR: MT1533.2; -		
DR	TubercuList: RV1488; -		
DR	InterPro: IPR001107; Band_7.		
DR	InterPro: IPR001972; Stomatin.		
DR	Pfam: PF01145; Band_7; 1.		
DR	SMART: SM00244; PHB-1.		
DR	PROSITE: PS01270; BAND_7; 1.		
TM	Hypothetical protein; Transmembrane; Complete proteome.		

FT	TRANSMEM	3	23	POTENTIAL
SQ	SEQUENCE	361 AA:	41282 MM:	32747A51D653CF91 CMC64:
	Query Match	26.7%:	Score 472:	DB 1; Length 381;
	Best Local Similarity	33.2%:	Pred. No. 1.2e-23:	
	Matches 110:	Conservative 72:	Mismatches 123:	Indels 26; Gaps 6:
QY	38	VLEPPOEAAMVVERMGFRHRLIEPGLNILLPYLDRIYVOSIKEIYINVEDOSAVTLDNV	97	
DB	26	VALIIOAEAAVIERLGRYSRTVSGLTFLPVIDRVARVDLREHVSFFPPVITEENL	85	
QY	98	TLQIDGVLYLRIMDPYKKSQYGVDEEYAVTOLAQTMSSELGKSLXDXVFRRESLNASI	157	
DB	86	TLNIDTVIEFOYTVQOAAVYEISNTIVGEGULTTLLRNVGGMLEOTLTRSRDINMQL	145	
QY	158	VDAINQADDCWGCIRCLRYEIKDHPHPVKEESQMOMQVAEERRKATVIESEGTRESAINV	217	
DB	146	REVLEDAIRGWRGLRARVARELRSIDPPISLOASMESQMOMADRRKAMILLTSEGTREAIKQ	205	
QY	218	AEKKQAOA-----ILASEEKAQINOAGASAVLAKAKAEAIRILAAALT	266	
DB	206	AEGOKQAOAILAEGAKQAKAAILLAEADROGRSMIRAOGEERAAAYLQACQAKAIEKTFPAI	264	
QY	267	QHNQGAASLTPAEQYVAFSEFKLAK-DSNFTI-TLPSNGDVTSMVAQAMGYG-----AL	319	
DB	265	--KACRPPEMIAIYQIYLOTLPEMAGMDKNNKVVVPSDFNALQGFTRLKGEDGVFRF	322	
QY	320	TKAPVPGTPDLSLSSGSSRDVQ-----TDAS	345	
DB	323	EPSYVEDQPKHNAADGDDAEVAWGWFSTDTDPDS	353	
RESULT	3			
YBBK	YBBK_ECOLI	STANDARD:	PRT:	305 AA.
AC	P7367:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein ybbk.			
GN	YBBK OR B0489 OR Z0642 OR ECS0552.			
OS	Escherichia coli and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN-K12 / MG1655.			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.,			
RT	"The complete genome sequence of Escherichia coli K-12."			
RL	Science 277:1453-1474(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Robertes D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,			
RA	Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kudl O.,			
RA	Lew H., Lin D., Namath A., Oetner P., Schramm S., Davis R.W.;			
RL	Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.			
RP	[3]			
RC	SEQUENCE FROM N.A.			
RX	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RX	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Postal G., Heckert J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocists K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Weich R.A., Blattner F.R.;			

```

RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." ;
RL      Nature 409:529-533(2001).
[4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / RIMD_0509952;
RX      MEDLINE=21156233; PubMed=11285796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
    Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M.,robe T.,
    Iida T., Takami H., Honda T., Sasakiwa G., Ogasawara N., Yasunaga T.,
    Kunara S., Shiba T., Hattori M., Shinagawa H.;
RA      RT
    "Complete genome sequence of enterohaemorrhagic Escherichia coli
    O157:H7 and genomic comparison with a laboratory strain K-12." ;
RL      DNA Res. 8:11-22(2001).
CC      -I SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between The Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@sib-sib.ch).
-----
CC      EMBL; AE000155; AAC73591.1; -
DR      EMBL; U82664; AAB40243.1; -
DR      EMBL; AE005230; AAG54846.1; -
DR      EMBL; AP002552; BAB33975.1; -
DR      EcoGene; EG13258; ybbK
DR      InterPro; IPR001107; Band_7.
DR      InterPro; IPR001972; Stomatin.
DR      Pfam; PF01145; Band_7; 1.
DR      SMART; SM00244; PHB; 1.
DR      PROSITE; PS01270; BAND_7; 1.
KW      Hypothetical protein; Transmembrane; Complete proteome.
FT      TRANSMEM 3 POTENTIAL.
SO      SEQUENCE 305 AA; 33743 MW; 0A7800F65790ACC3 CRC64;

Query Match          26.4%; Score 466; DB 1; Length 305;
Best Local Similarity 34.8%; Pred. No. 2,2e+23;
Matches 98; Conservative 76; Mismatches 94; Indels 14; Gaps 6;

QY      38 VLFVPOQLAAWYVERGRHRRLLEPELNLIPYLDIRIVOSIKELVINVPEDSAVTLDNV 97
        ||| | | | | | | | | : | | | | | | | | | : | | | | | : | |
DB      20 VKPIPGQGYMTVERGRRTKTLPGLSLVFPMDFIRGRINMMEDVLDPISQEVISKDNA 79

QY      98 TLQIDGVLYLRIMDPYKASGYDEPDEVTOQAORTMSESGKLSDXRVFPRESLNASTI 157
        || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      80 NVITDAVCFIQIDAPRAYEVSNELAIINLTMINIRTVLGSMDELDSORDINSGL 139

QY      158 VDAINAADCGICRCLRYEKIDIHVPVKVESMOQVVEARRKATVLESGETRESAINV 217
        : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
DB      140 LRIVDEATNPNGICKYTRIEIRDVRPAELLISMNQMAEFRTKRAYIIILAEIGROAEILK 199

QY      218 AEGKKAOQAILLSEAEKAQINOA--GEASAVYLAKAKAEALRIILAALTQNCD--AAAS 275
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      200 AGEKEQSOLLAERGGSAFLQAEERKS----AEAEKARKMSSEAIA--SGDIQAVN 252

QY      276 LTVAEOVYSASFSLAKDSENT--ILLPSNPGDYTSNVAQAMGV 315
        ||::|| | | | | | | | | | | | | | | | | | | | | | | |
DB      253 YFAOKYTEALQIGSSSNRSKVVMPL--EASSLMGSIAGI 291

RESULT      4
ID          Y692_METTH STANDARD; PRT; 318 AA.
AC          O26788;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical protein MTH692.
OS      Methanobacterium thermoautotrophicum.
```

OC	Archaea:uryarchoaeota; Methanobacteriales; Methanobacteriaceae;
OC	Methanothermobacter.
OX	NCBI_taxid=145262;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=DELTA H;
FX	MEDLINE=98037514, PubMed=9371463;
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubots J.,
RA	Alredede T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA	Spadafora R., Violare R., Wang Y., Merzowski J., Gibson R.,
RA	Jivani N., Carno A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA	McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA	Danels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT	"Complete genome sequence of <i>Methanobacterium thermoautotrophicum</i>
RT	deltah: functional analysis and comparative genomics.";
RL	J. Bacteriol. 179:7135-7155(1997).
CC	-I- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage of it and for commercial
CC	entitles requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: AE000848; AAB85197.1; -
DR	InterPro: IPR001107; Band_7; -
DR	InterPro: IPR001972; Stomatin.
DR	Pfam: PF01145; Band_7; 1.
DR	PRINTS: PR00721; STOMATIN.
DR	SMART: SM00244; PHB; 1.
DR	PROSITE: PS01270; BAND_7; 1.
KW	Hypothetical protein; Transmembrane; Complete proteome.
FT	TRANSMEM 2
SO	SEQUENCE 318 AA; 35425 MW; 44720CAC8E99278D CRC64;
	POTENTIAL.

[illegible]

OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 CC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A. 4304 / ATCC 49558;
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 Reichenbach K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 Richardson D.L., Kierkegaard A.R., Graham D.E., Kyriides N.C.,
 Fleischman R.D., Overbeek R., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 Kikunas E.F., Retch C.I., McNeill L.K., Badger J.H., Glodex A., Zhou L.,
 Peterson S., Goeyne J.D., Weidman J.F., McDonald L., Uterback T.,
 Colton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RT Nature 390:364-370(1997).
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE001006; AAB89829.1; -
 DR TIGR: AF1420; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatoln.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE: PS01270; BAND_7; 1.
 DR Hypothetical protein: Transmembrane; Complete proteome.
 KW TRANSMEM 3 23 POTENTIAL.
 FT SEQUENCE 249 AA; 28346 MW; 4AF8EAD58AF6F230 CAC64;
 SQ
 Query Match 20.5%; Score 362; DB 1; Length 249;
 Best Local Similarity 36.0%; Pred. No. 8.5e-17;
 Matches 81; Conservative 50; Mismatches 76; Indels 18; Gaps 5;
 Oy 37 VVLF-----VPOEAVNVERMGRFRHLEPGLNLIPLVDLRIRYVQSLKEIYINVPESQ 89
 Db 12 VVFLLSAVRIKVEKRGVIFRLGVLGARGGLFIIPILENM-VVVDLRITVDVPSQ 70
 Oy 90 SAVLTNTLTQIDGVLYLRIMDPYKASYGVEDPEYAVTQLOQTMRSELGKLSXDRVRESLNASI 149
 Db 71 EVYTDNTVAVNAVYRVVDPKAAVTEVPYQYATQALOQTIRSLIIGQELDEVLSLSE 130
 Oy 150 RESLNASTVDAINQADWCIGRLCLREIKDIHVPPRVKESQMOMQVEARERKRAVLESEGR 209
 Db 131 RDKLVKRLQOITIDEETNPMGIVKVTAVEIKDVELPEEMRRIMQAQAEERERSKILIRAE 190
 Oy 210 TRESAINVAEGKKOQILASEAKA-----EQINQAGASAVL 248
 Db 191 EYQAAMKRLREA--ADVLA-OSEGAILRLRYLTQTLNEISAEQNTTI 231
 RESULT 6
 ID Y827_METUA STANDARD; PRT; 199 AA.
 AC Q58237;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein M20827.
 GN M20827.
 OS Methanococcus jannaschii.
 CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Goeyne J.D.,
 Kierkegaard A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 Overbeek R., Kikunas E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
 Scott J.L., Geoghegan N.S.M., Weidman J.F., Fulmerman J.L., Nguyen D.,
 Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U67526; AAB98826.1; -
 DR TIGR: M20827; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatoln.
 DR Pfam: PF01145; Band_7; 1.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE: PS01270; BAND_7; 1.
 DR Hypothetical protein: Transmembrane; Complete proteome.
 KW TRANSMEM 7 27 POTENTIAL.
 FT SEQUENCE 199 AA; 22809 MW; 2545CE07DD54337E CAC64;
 SQ
 Query Match 20.1%; Score 354.5; DB 1; Length 199;
 Best Local Similarity 38.5%; Pred. No. 2e-16;
 Matches 67; Conservative 46; Mismatches 60; Indels 1; Gaps 1;
 Oy 38 VLFVPOEAVNVERMGRFRHLEPGLNLIPLVDLRIRYVQSLKEIYINVPESQAVTLDNV 97
 Db 25 IVIVNQEGLIFRLGVLGARGGLFIIPILENM-VVVDLRITVDVPSQ 70
 Oy 98 TQIDGVLYLRIMDPYKASYGVEDPEYAVTQLOQTMRSELGKLSXDRVRESLNASI 157
 Db 84 VYKVDVAVYVRIYDEKALILEVEDYEVAVINLAQTTLAIGSMELDEVLNKREYNKSL 143
 Oy 158 VDAINQADWCIGRLCLREIKDIHVPPRVKESQMOMQVEARERKRAVLESEGR 211
 Db 144 LEILDRETDAMGVRIEKEVKEVIDPEIDIKNMAQMAERLRAALILEAEGER 197
 RESULT 7
 ID YF11_PYRHO STANDARD; PRT; 266 AA.
 AC O59180;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein PH1511.
 GN PH1511.
 OS Pyrococcus horikoshii.
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]


```

RP SEQUENCE FROM N.A.
RC STRAIN-OT3:
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Hatakeyama Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtoku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -i SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AP000006; BAA30619.1; -
CC InterPro: IPR001107; Band_7.
CC InterPro: IPR001972; Stomatlin.
CC Pfam: PF01145; Band_7; 1.
CC PRINTS: PR00721; STOMATLIN.
CC SMART: SM00244; PHB; 1.
CC PROSITE: PS01270; BAND_7; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
CC TRANSMEM 7
CC SEQUENCE 266 AA; 29999 MW; 21D06F199C31021F CRC64;
SO

Query Match 18.5%; Score 327.5; DB 1; Length 266;
Best Local Similarity 33.0%; Pred. No. 1.5e-14;
Matches 72; Conservative 55; Mismatches 82; Indels 9; Gaps 3;

OY 36 TVLVLPPOEAWVEMRGFRHLEPGILNIPVDRIRYVSLKEIVNPEQSAVTD 95
DB 23 SAIKIVKEKRAVIRLGRVVGARGPGLFIIPIFEK-AVYIDLKRVQVDVYQETITD 81
OY 96 NVTLQIDGLYLRLMDPYKASYGEDEPEYAVTQLAQTMRSELGKSLDXFRERESLNA 155
DB 82 NVPVHNNAVYFRVVDVPAVAVQVKNYIMATQSITOTLRSVIGQAHDELLESLERDKLNM 141
OY 156 STVDAINQADWCGRCLRYEIKDTHVPRVKEQMOYVEARRKRAVLESGRRESAI 215
DB 142 QLQRIIDETDPMGKIVYAVEIKDELPLPGMQRAMARQAERERRRRTITLAEARQAAE 201
OY 216 NVAEGKKOAOI-----LASEAKAEQINOAGEASAVL 248
DB 202 KUREA---AEIIEHPMALQLRTLOTISDVADKSNVI 236

```

```

RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
CC -i SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AJ248285; CAB9571.1; -
CC InterPro: IPR001107; Band_7.
CC InterPro: IPR001972; Stomatlin.
CC Pfam: PF01145; Band_7; 1.
CC PRINTS: PR00721; STOMATLIN.
CC SMART: SM00244; PHB; 1.
CC PROSITE: PS01270; BAND_7; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
CC TRANSMEM 1
CC SEQUENCE 268 AA; 30241 MW; AELFBBBD48A4F84 CRC64;
SO

Query Match 18.2%; Score 321.5; DB 1; Length 268;
Best Local Similarity 31.9%; Pred. No. 3.8e-14;
Matches 76; Conservative 54; Mismatches 83; Indels 25; Gaps 5;

OY 32 LPRN-----TVVL-----TVPOEAWVEMRGFRHLEPGILNIPVDRIRY 75
DB 3 LPTNFFVTIILFIIFLTAIAKIVKEKRAVIRLGRVVGARGPGLFIIPIFEK-AV 61
OY 76 VQSLKEIVNPEQSAVTLNDVTLQIDGLYLRLMDPYKASYGEDEPEYAVTQLAQTMR 135
DB 62 IYDLKRVQVDVYQETITDNDVPRVNAVYFRVVDVPAVAVQVKNYIMATQSITOTLR 121
OY 136 SELGKSLDXFRERESLNAIVDAINQADWCGRCLRYEIKDTHVPRVKEQMOYVE 195
DB 122 SYVQAHDELLESLERDKLNMQRIIDEATDPMGKIVYAVEIKDELPLPGMQRAMAKQAE 181
OY 196 AERRKRAVLESGRRESAINVAEGKKOAOI-----LASEAKAEQINOAGEASAVL 248
DB 182 AERRRARTITLAEARQAAEKUREA---AEIIEHPMALQLRTLOTISDVADKSNVI 236

```

```

RESULT 9
BAN7_HUMAN STANDARD; PRT; 287 AA.
AC P27105; Q14087; Q15609;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythrocyte band 7 integral membrane protein (Stomatlin) (Protein
DE 7.2b).
GN EPB72 OR BMD7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91355220; PubMed=183838;
RA Hiebl-Dierschmid C.M., Entler B., Glotzmann C., Maurer-Fogy I.,
RA Stratowa C., Prohaska R.;
RT "Cloning and nucleotide sequence of cDNA encoding human erythrocyte
RT band 7 integral membrane protein.";
RL Biochim. Biophys. Acta 1090:123-124(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96423038; PubMed=8825639;
RA Unfried I., Entler B., Prohaska R.;
RT "The organization of the gene (EPB72) encoding the human erythrocyte
RT band 7 integral membrane protein (protein 7.2b).";
RL Genomics 30:521-528(1995).

```

```

RA RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Stewart A.C., Hepworth-Jones B.J., Keen J.N., Dash B.J.C.,
RA Argent A.C., Casimir C.M.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=96064711; PubMed=7592848;
RA Gallagher P.G., Forget B.G.;
RT "Structure, organization, and expression of the human band 7.2b gene,
RT a candidate gene for hereditary hydrocytosis.";
RL J. Biol. Chem. 270:26358-26363(1995).
RN
RP SEQUENCE OF 4-24, AND PHOSPHORYLATION SITE.
RX MEDLINE=93385136; PubMed=8373790;
RA Salzer U., Aorn H., Prohaska R.;
RT "Identification of the phosphorylation site on human erythrocyte band
RT 7 integral membrane protein: implications for a monotypic protein
RT structure.";
RL Blochlm. Biophys. Acta 1151:149-152(1993).
RN
RL
RN PALMITOYLATION.
RX MEDLINE=99268434; PubMed=10386112;
RA Snyers L., Umlauf E., Prohaska R.;
RT "Cysteine 29 is the major palmitoylation site on stomatin.";
RL FEBS Lett. 449:101-104(1999)
CC
CC -1- FUNCTION: THOUGHPT TO REGULATE CATION CONDUCTANCE.
CC -1- SUBCELLULAR LOCATION: EXPOSED ON THE CYTOPLASMIC SURFACE OF THE
CC MEMBRANE.
CC -1- DISEASE: DEFECTS IN EPB72 IN RED CELLS OF PATIENTS WITH HEREDITARY
CC STOMATOCYTOSIS OR CRYOHYDROCYTOSIS RESULT IN AN INCREASED NA+/K+-
CC PERMEABILITY AND HENCE TO A DISORDER OF CELL VOLUME CONTROL.
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X60067; CAA42671.1; -.
DR EMBL; X85116; CAA59436.1; -.
DR EMBL; X85117; CAA59436.1; JOINED.
DR EMBL; M81635; AAA58432.1; -.
DR EMBL; U33931; AAC50296.1; ALT-SEQ.
DR EMBL; U33925; AAC50296.1; JOINED.
DR EMBL; U33926; AAC50296.1; JOINED.
DR EMBL; U33927; AAC50296.1; JOINED.
DR EMBL; U33928; AAC50296.1; JOINED.
DR EMBL; U33929; AAC50296.1; JOINED.
DR EMBL; U33930; AAC50296.1; JOINED.
DR PIR; S17659; S17659.
DR MIM; 133090; -.
DR MIM; 185000; -.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS; PRO0721; STOMATIN.
DR SMART; SM00244; PHB; 1.
DR PROSITE; PS01270; BAND_7; 1.
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.
FT INT_MPT 0
FT TRANSMEM 25 53 POTENTIAL.
FT DOMAIN 54 287 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 9 9 PHOSPHORYLATION (BY CAPK).
FT LIPID 29 29 PALMITATE.
FT LIPID 86 86 PALMITATE (PARTIAL).
FT CONFLICT 5 5 H -> D (IN REF. 3)
SQ SEQUENCE 287 AA: 31599 MW: 4852FE286ECC337 CRC64:

```

```
Query Match          15.9% ; Score 281.5; DB 1; Length 287;  
Best Local Similarity 30.0%; Pred. No. 1,6e-11;  
Matches    65; Conservative   52; Mismatches   77; Indels   23; Gaps     5
```

OY 36 TVVLF-----VPQGEAVWVERMGFRHILE-----PGNLILPYLVDRIKYQSLLK 80
 |||::||
Db 41 TWTFTPIISMCKIKIKEERAILPRLG---RLLGGAGPKPLFLFLLPTDSFKIV--DDR 96

OY 81 ETIVNWPEOSATVLNVTLOIDGVLYLRIMDPDYKASGYEDEVAAYTOLACTMTSELGR 140
 :||::||::||||::|:
Db 97 TISPDIPQEELITKRGSVTSYDGVYYIRQNQTMLVANNTMDSATRLLAQTLTRNVLT 156

OY 141 LSXDKVFRESLSNIYDAINQAADCWICIGCLRREIKDIHVPRVKESMQNQVEAEERRK 200
 :::||::|::|::|::|::|::|:
Db 157 KKLSQLSRSREIAHNMOSTLDADWDAMCIKEREAIKVDCIPLVOLQRMAAAEAESAAREA 216

OY 201 RATVLESEGTREGAINVAEGKKOAOTILAASEAKAEOI 237
 |||::|:
Db 217 RAKTVTAABG----EMNRSAALKAEASVVITESPAALQL 249

[illegible]

```

Db 114 VSVAHVVFRTSDPIASVNNVDADIVSTKLLAQPTTLRNALGMKLTUTEMLTREARLQJCE 173
Oy 159 DAINQADCMKRIICRLRYIKINIHVPPRKESMQVQAEERRKATVLESECTRES--AIN 216
Db 174 TILLEGTHMVKVKEVREVKIDIRLPQOULTRAMAEEAERAKRKYVAAGEQKASRLK 233
Oy 217 VAEGKQAOILASEAEKAEQINOAGASAVL 248
Db 234 EADVYQANPVALQRLQALNSTIAAEHNSTI 265

RESULT 11
BAN7_MOUSE
ID BAN7_MOUSE STANDARD: PRT: 284 AA.
AC PS4116; G60744; Q62455.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Erythrocyte band 7 integral membrane protein (Stomatlin) (Protein
DE 7.2B)
CN EPB72 OR EPB7.2.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A.
RP STRAIN-BALB/C; TISSUE=Bone marrow;
RX MEDLINE=97080556; PubMed=8921901;
RA Stiegel W., Unfried J., Prohaska R.;
RT "Cloning and analysis of a cDNA encoding the BALB/c murine
RT erythrocyte band 7 integral membrane protein.";
RL Gene 178:115-118(1996).
[2]
RN RP SEQUENCE FROM N.A.
RP STRAIN-C57BL/6J;
RA Gallagher P.G., Romana M., Ileana J.H., Ward D.C.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RP MEDLINE=96374834; PubMed=8786142;
RA Gallagher P.G., Turetsky T., Mentzer W.C.;
RT "Genomic organisation and 5'-flanking DNA sequence of the murine
RT stomatin gene (EpB72).";
RL Genomics 34:410-412(1996).
-1- FUNCTION: THOUGHT TO REGULATE CATION CONDUCTANCE.
-1- SUBCELLULAR LOCATION: EXPOSED ON THE CYTOPLASMIC SURFACE OF THE
MEMBRANE.
-1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstati-
CC c the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for commen-
CC c entities requires a license agreement (See http://www.isb-sib.ch/anno
CC or send an email to license@isb-sib.ch).
-----
Cc CC
Cc EMBL; X91043; CAAG5203.1; -
DR EMBL; U17297; AAA75024.1; -
DR EMBL; U50999; AAB18857.1; ALT_SEQ.
DR EMBL; U50993; AAB18857.1; JOINED.
DR EMBL; U50994; AAB18857.1; JOINED.
DR EMBL; U50995; AAB18857.1; JOINED.
DR EMBL; U50996; AAB18857.1; JOINED.
DR EMBL; U50997; AAB18857.1; JOINED.
DR EMBL; U50998; AAB18857.1; JOINED.
DR MCD; MG1.95403; EPB7.2.
DR InterPro: IPR001107; Band.7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band.7; 1.
DR PRINTS; PRO0721; STOMATIN.
DR SMART; SM00244; PHB.1.

```

Query Match	Best Local Similarity	15.7%; Score 277.5; DB 1; Length 284;
Matches	63; Conservative	50; Mismatches 79; Indels 13; Gaps 4;
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		N -> H (IN REF. 1).
SQ SEQUENCE 284 AA; 31403 MW; 8E49276611ACB84F CMC64;		
DR PROSITE PS01270; BAND_7; 1.		
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.		
FT TRANSMEM 32 52		POTENTIAL.
FT DOMAIN 53 284		CYTOPLASMIC (POTENTIAL).
FT LIPID 30 30		PALMITATE (BY SIMILARITY).
FT LIPID 87 87		PALMITATE (PARTIAL) (BY SIMILARITY).
FT CONFLICT 37 37		V -> A (IN REF. 2).
FT CONFLICT 40 40		I -> F (IN REF. 2).
FT CONFLICT 43 43		L -> I (IN REF. 2).
FT CONFLICT 91 91		F -> L (IN REF. 2).
FT CONFLICT 273 273		V -> I (IN REF. 1).
FT CONFLICT 283 283		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:20:57 ; Search time 63.67 Seconds

(without alignments)
896.843 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767
Sequence: 1 MLARAARGHMGPPAEGLSTG.....RDVQGTDAKXDEELDRVKMS 356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646	93.2	356	4	Q9UJZ1
2	1639	92.8	356	4	Q96FY2
3	1613	91.3	342	4	Q9P042
4	1605	90.8	357	4	Q60376
5	1537.5	87.0	353	11	Q99JB2
6	1526.5	86.4	353	11	Q9DCG8
7	1064.5	60.2	323	5	Q9WIF7
8	957.5	54.2	334	5	Q9XVP9
9	861.5	48.8	354	3	Q60121
10	826	46.7	411	10	Q93VP9
11	826	46.7	515	10	Q9V082
12	816.5	46.2	401	10	Q9LW00
13	811.5	45.9	312	16	Q92IC8
14	806	45.6	311	16	Q9ZDK0
15	794.5	45.0	394	10	Q9M585
16	722	40.9	374	5	O77376

17	716	40.5	315	16	Q9JR80	Q9JR80 neisseria m
18	714.5	40.4	315	2	Q9JPH5	Q9JPH5 neisseria m
19	714	40.4	315	2	Q9JPN3	Q9JPN3 neisseria m
20	712.5	40.3	315	2	Q9JPC5	Q9JPC5 neisseria m
21	712.5	40.3	315	16	Q9JRP6	Q9JRP6 neisseria m
22	711.5	40.3	315	2	Q9K2T7	Q9K2T7 neisseria g
23	711.5	40.3	315	2	Q9JPH3	Q9JPH3 neisseria m
24	711.5	40.3	315	2	Q9JPN2	Q9JPN2 neisseria m
25	664	37.6	357	5	Q9U4M5	Q9U4M5 leishmania
26	497	28.1	344	16	Q9RS11	Q9RS11 delinococcus
27	496.5	28.1	337	16	Q9PGK3	Q9PGK3 xylolla fas
28	486	27.5	310	16	Q9A331	Q9A331 caudobacter
29	477	27.0	306	2	Q9AH28	Q9AH28 photorhabd
30	475	26.9	319	2	Q9X9Z6	Q9X9Z6 streptomyce
31	474.5	26.9	325	2	Q9F507	Q9F507 escherichia
32	473	26.8	307	16	Q9CMB1	Q9CMB1 pasteurella
33	467.5	26.5	306	16	Q9KTC6	Q9KTC6 vibrio chol
34	464.5	26.3	343	2	Q9K458	Q9K458 streptomyce
35	460	26.0	298	17	Q98205	Q98205 pyrococcus
36	460	26.0	299	17	Q9UYE4	Q9UYE4 pyrococcus
37	454	25.7	328	16	Q9ZM51	Q9ZM51 rhizobium l
38	444.5	25.2	374	16	Q9CBM5	Q9CBM5 mycobacteri
39	442.5	25.0	322	16	Q97K67	Q97K67 clostridium
40	442	25.0	320	2	Q9R0U5	Q9R0U5 clostridium
41	430	24.3	316	16	Q98H59	Q98H59 rhizobium l
42	423.5	24.0	305	16	Q9WZK0	Q9WZK0 thermotoga
43	417.5	23.6	392	17	Q9HSA8	Q9HSA8 halobacteri
44	412	23.3	309	16	Q98RE1	Q98RE1 mycoplasma
45	376.5	21.3	253	16	Q67057	Q67057 aquilex aeo

ALIGNMENTS

RESULT 1	Q9UJZ1	PRELIMINARY:	PRT:	356 AA.
ID	Q9UJZ1			
AC	Q9UJZ1			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	MEMBRANE ASSOCIATED PROTEIN SLP-2 (STOMATIN-LIKE PROTEIN 2) (STOMATIN-LIKE 2) (CDNA FLJ14499 FIS, CLONE NT2RMI00080, WEAKLY SIMILAR TO HUSLP2 OR SLP2).			
GN	HUSLP2 OR SLP2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
NCBI_TaxID	9606;			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=HEART MUSCLE;			
RA	Wang Y., Morrow J.S.;			
RT	"Identification and Characterization of Human SLP-2, a Novel Homologue of Stomatlin (band 7.2b) Present in Erythrocytes and Other Tissues.";			
RT	Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=BRAIN;			
RC	Owczarek C.M., Treutlein H.R., Portbury K.J., Gulluyan L.M., Kola I., Hertzog P.J.;			
RA	"A novel member of the stomatin/EPB2/MC-2 family, stomatin-like protein 2 (SLP2), is ubiquitously expressed and localizes to HSA chromosome 9p13.1.";			
RT	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=LUNG CARCINOMA, SKIN, AND MELANOTIC MELANOMA.;			
RA	Strusberg R.;			
RL	Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Isoqal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			

RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
 RA Megatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Makematsu A., Nakamura Y., Nagahari K., Masuo Y.,
 RA Niinomiya K., Iwyanagi T.,
 RT "NEDO human cDNA sequencing project."
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS, EPITHELIOID CARCINOMA;
 RA Strausberg R.,
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF190167; AAF09142.1; -
 DR EMBL: AF282596; AAF91466.1; -
 DR EMBL: BC003025; AAH03025.1; -
 DR EMBL: BC002442; AAH02442.1; -
 DR EMBL: AK027405; BAB55081.1; -
 DR EMBL: BC014990; AAH14990.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatlin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATLIN.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 356 AA; 38534 MW; 672331B57C82654E CRC64;

Query Match 93.2%; Score 1646; DB 4; Length 356;
 Best Local Similarity 95.3%; Pred. No. 1.2e-97;
 Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

OY 1 MLARAARGHMPFAEG--LSTGFMPRSGRASSGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
 DB 1 MLARAARGTGCALLLRLGSLASGRAPR--RASSGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
 OY 59 LEPLGNTLIPVLDRIYQSLKEIYINPEQSAVTLDNVTLQIDGVLRLIMDPYKASYG 118
 DB 59 LEPLGNTLIPVLDRIYQSLKEIYINPEQSAVTLDNVTLQIDGVLRLIMDPYKASYG 118
 OY 119 VEDPEYAVTQLAQTMRSELKSLDKYFRRESINASTIVAINOACMCIRCLRYEIK 178
 DB 119 VEDPEYAVTQLAQTMRSELKSLDKYFRRESINASTIVAINOACMCIRCLRYEIK 178
 OY 179 DIHVPVRKESMOMQVEARRKRATVLESEGTRESAINVAEGKKQAQILASEAEKAEQIN 238
 DB 179 DIHVPVRKESMOMQVEARRKRATVLESEGTRESAINVAEGKKQAQILASEAEKAEQIN 238
 OY 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEOYVSASFSLAKDSNTILL 298
 DB 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEOYVSASFSLAKDSNTILL 298
 OY 299 PSNPGDVTSMVAQAMGVGALTAKAPVPGTSPSLSSGSRDVGDTASLDEELDRYKMS 356
 DB 299 PSNPGDVTSMVAQAMGVGALTAKAPVPGTSPSLSSGSRDVGDTASLDEELDRYKMS 356
 RESULT 2
 O96FY2 PRELIMINARY; PRT; 356 AA.
 AC O96FY2
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE STOMATLIN-LIKE 2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.,
 RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC010152; AAH10152.1; -
 SQ SEQUENCE 356 AA; 38518 MW; 6FC0BF899E1919A7 CRC64;
 Query Match 92.8%; Score 1639; DB 4; Length 356;
 Best Local Similarity 95.0%; Pred. No. 3.4e-97;
 Matches 340; Conservative 2; Mismatches 12; Indels 4; Gaps 2;

OY 1 MLARAARGHMPFAEG--LSTGFMPRSGRASSGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
 DB 1 MLARAARGTGCALLLRLGSLASGRAPR--RASSGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
 OY 59 LEPLGNTLIPVLDRIYQSLKEIYINPEQSAVTLDNVTLQIDGVLRLIMDPYKASYG 118
 DB 59 LEPLGNTLIPVLDRIYQSLKEIYINPEQSAVTLDNVTLQIDGVLRLIMDPYKASYG 118
 OY 119 VEDPEYAVTQLAQTMRSELKSLDKYFRRESINASTIVAINOACMCIRCLRYEIK 178
 DB 119 VEDPEYAVTQLAQTMRSELKSLDKYFRRESINASTIVAINOACMCIRCLRYEIK 178
 OY 179 DIHVPVRKESMOMQVEARRKRATVLESEGTRESAINVAEGKKQAQILASEAEKAEQIN 238
 DB 179 DIHVPVRKESMOMQVEARRKRATVLESEGTRESAINVAEGKKQAQILASEAEKAEQIN 238
 OY 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEOYVSASFSLAKDSNTILL 298
 DB 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEOYVSASFSLAKDSNTILL 298
 OY 299 PSNPGDVTSMVAQAMGVGALTAKAPVPGTSPSLSSGSRDVGDTASLDEELDRYKMS 356
 DB 299 PSNPGDVTSMVAQAMGVGALTAKAPVPGTSPSLSSGSRDVGDTASLDEELDRYKMS 356

RESULT 3
 O9P042 PRELIMINARY; PRT; 342 AA.
 AC O9P042
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HSPC108.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,
 RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.J.,
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 300 previously undefined genes expressed in CD34+ hematopoietic
 stem/progenitor cells."
 RL Genome Res. 10:1546-1560(2000).
 DR EMBL: AF161458; AAF29073.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatlin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATLIN.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 342 AA; 37145 MW; 7958C0E3BDBE53E5 CRC64;

Query Match 91.3%; Score 1613.5; DB 4; Length 342;
 Best Local Similarity 97.1%; Pred. No. 1.4e-95;
 Matches 330; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
 OY 17 LSTGFMPRSGRASSGLPRNTVLFVPOQEAAMVVERMGFRHRILEPLGNTLIPVLDRIYV 76
 DB 6 LASC---RACRASSGLPRNTVLFVPOQEAAMVVERMGFRHRILEPLGNTLIPVLDRIYV 62

Oy	77	OSLKEIYIVNEQOSAVTLNDNTLTLOIDGLYLRIMDPKASGVNDEPREAYTOLAQTMR	136
Db	63	QSLKEIYIVNEQOSAVTLNDNTLTLOIDGLYLRIMDPKASGVNDEPREAYTOLAQTMR	122
Oy	137	ELGKLSXKVERERESLNASTVDAINOAADCGICRLRYEIKDIHBPVYKESMOQVEA	196
Db	123	ELGKLSLTDKVRERESLNASTVDAINOAADCGICRLRYEIKDIHBPVYKESMOQVEA	182
Oy	197	ERRRPAVLEEEGRRESAINVAEKKQAQLLASAEKAEQINOAGGASAVLAKAKAAE	256
Db	183	ERRRPAVLEEEGRRESAINVAEKKQAQLLASAEKAEQINOAGGASAVLAKAKAAE	242
Oy	257	AIRILAAALTOHNDAAASLTVAEOYSAESKLAKDSNTLTLPSPEDVYSMTAQAAMGY	316
Db	243	AIRILAAALTOHNDAAASLTVAEQYSAESKLAKDSNTLTLPSPEDVYSMTAQAAMGY	302
Oy	317	GALTAPVPGTSPDLSLSSSSNDVOGTPASXADELDRKMS	356
Db	303	GALTAPVPGTSPDLSLSSSSNDVOGTPASLDEELDRKMS	342

RESULT	4	
060376		
ID	060376	PRELIMINARY;
AC	060376.	PRT; 357 AA.
DT	01-AUG-1998	(TREMBLrel. 07, Created)
DT	01-AUG-1998	(TREMBLrel. 07, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE-FIBROBLAST OR FORESKIN;
RA Lamerich J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stllwagen S.,
RA Phan H., Velasco N., Ganes J., Dangman L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Adreaise T.,
RA Trankhelm M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Ariellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,

RT "Sequence analysis of a human P1 clone containing the XRCC9 DNA repair
RT gene.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

SQ SEQUENCE 357 AA; 38749 MW; 512632B83028135A CRC64;

Query Match	90.8%	Score 1605	DB 4	Length 357
Best Local Similarity	93.8%	Pred. No. 5.1e95		
Matches 332	Conservative 2	Mismatches 4	Indels 16	Gaps 2

```

0Y 17 L$GFNFRSGA$GLPRNRYVLFVPOQEA$VYERMRFRHLEPGNILLIPVLDRI$RY 76
    ! ! ! ! !
Db 6 L$SGR$P - - R$SGGLPRNRYVLFVPOQEA$VYERMRFRHLEPGNILLIPVLDRI$RY 63

0Y 77 Q$LKEIYINPEQ$AVLNDNTLOIDGV$LIRIMDPYK$GVEDPEYAVTOLAQ$TMS 136
    ! ! ! ! !
Db 64 Q$LKEIYINPEQ$AVLNDNTLOIDGV$LIRIMDPYK$GVEDPEYAVTOLAQ$TMS 123

0Y 137 ELGK$XDKYFRERES$IN$IVDAINQ$ADCMGIRCLRYEIKDIHVPYK$ESM$Q - - 193
    ! ! ! ! !
Db 124 ELGK$LDK$YFRERES$IN$IVDAINQ$ADCMGIRCLRYEIKDIHVPYK$ESM$QVGA 183

0Y 194 -----VE$ER$RR$RATVLE$E$E$TRES$ATIN$VAGK$KQAO$IL$SE$AK$AO$INQ$AG 242
    ! ! ! ! !
Db 184 KEGWEKGL$R$AV$E$ERR$RATVLE$E$E$TRES$ATIN$VAGK$KQAO$IL$SE$AK$AO$INQ$AG 243

```

QY	243	ESASVLAFAKKAKKAEIRILTAALTOHNDGDAASLTVEAOYVSASKLAKOSNTLLPSNP	302
Db	244	ESASVLAFAKKAKKAEIRILTAALTOHNDGDAASLTVEAOYVSASKLAKOSNTLLPSNP	303
QY	303	GDVYSMAVAQAMGVGALTKAPFPGTPOSLSSGSSRDVQGTDAKDELDTRYKMS	356
Db	304	GDVYSMAVAQAMGVGALTKAPFPGTPOSLSSGSSRDVQGTDAKDELDTRYKMS	357

RESULT 5
Q99JB2
ID Q99JB2 PRELIMINARY; PRT; 353 AA

DT 01-JUN-2001 (TREMBLrel. 17, created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE RIKEN CDNA 0610038F01 GENE (STOMATIN-LIKE PROTEIN 2)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases

RP SEQUENCE FROM N.A.
RC STRAIN-FVB; TISSUE=KIDNEY;
RA Chang J.G., Chan W.L.;
RT "Mouse stomatin-like protein 2 (MSLP2) mRNA."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases

DR EMBL: BC003425; AAH03425.1; -
DR EMBL: AF323178; AAG53404.1; -
DR MGD: MGI:1913842; 0610038F01.RIK.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatn.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 353 AA; 38385 MW; 391D269576F6E6B8 CRC64;

Query Match	87.0%;	Score 1537.5;	DB 11;	Length 353;
Best Local Similarity	89.9%;	Pred. No. 1e-90;		
Matches 322;	Conservative 8;	Mismatches 21;	Indels 7;	Gaps 4

1 MEANNNKNS I GALEEKVS YQASGRVFR - KASSGELFKNI VILF VFQQEAM

59 LERGLNVLIPVLDLRIRYVQSLNEIVINVEQSAVTLIDNVTLQIDGVLYLRIMDPYKASYG 118
119 VEDPEYAVTQLAQTTMRSELGKISXDKYPERERESLNASIVDAINQAAQCGIRCLYEYIK 178

119 VEDPEAVIYQLOATOTYRSELGKLSLDKVPREESIINAIIVAINOMADCWICRCLPREIK 178b
179 DIIVPRVKESMOMVEARRRRATVLSSEGRSAIVAGCKQAOQLSLEAKAQIN 238y

239 QAGEASAVLAKAKAKAEAIRILAAALWQHNGDAAASLTVAEQVVSASFSLAKDSNTLL 298
|||||
239 QAGEASAVLAKAKAKAEAIRILAGALTOHNGDAAASLTVAEQVVSASFSLAKDSNTVLL 298
|||||

b

```
299 PSNPEDVTSMVAQAMGVYGA LTKAPVPPTDLS SSGSSRDVQGTASXDEELDRVKMS 356  
||||| ||||| : ||| ||| ||| |||  
299 PSNPDSVTSMVAQAMGVYGA LTKAPVPGAQN -SSQSRDVGATPSTI-EELGRVKLS 353
```

```
RESULT 6
ID Q9DCG8 PRELIMINARY: PRT: 353 AA.
AC Q9DCG8:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 0610038F01RIK PROTEIN.
GN 0610038F01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamija M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Notone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL: AK002793; BAB2363.1; -;
DR MGD: MGI:1913842; 0610038F01RIK.
DR InterPro: IPR001107; Band.7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band.7; 1.
DR PRINTS: PR00721; STOMATLIN.
DR SMART: SM00244; PHB; 1.
SO SEQUENCE 353 AA; 38530 MW; D317CBB7E32F8863 CRC64;

Query Match 86.4%; Score 1526.5; DB 11; Length 353;
Best Local Similarity 89.4%; Pred. No. 5,2e-90;
Matches 320; Conservative 8; Mismatches 23; Indels 7; Gaps 4;
```

```
DB 299 PNPSPDVTSMVAQAMGVGALTKAPVPGAON--SSQSRDVOATDTST-EEELGRVYKLS 353
RESULT 7
ID Q9WLF7 PRELIMINARY: PRT: 323 AA.
AC Q9WLF7:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG2970 PROTEIN.
GN CG2970.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borrova D., Botchan M.R., Bouck H., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski R., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner C., Turner E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of Drosophila melanogaster.*;
RL Science 287:2185-2195(2000).
DR EMBL: AE003462; AAF47110.1; -;
DR FlyBase: FBgn0034936; CG2970.
DR InterPro: IPR001107; Band.7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band.7; 1.
DR PRINTS: PR00721; STOMATLIN.
DR SMART: SM00244; PHB; 1.
SO SEQUENCE 323 AA; 35668 MW; D5C6241445F8F4DE CRC64;
```

```
Query Match 60.2%; Score 1064.5; DB 5; Length 323;
Best Local Similarity 63.7%; Pred. No. 1,6e-60;
Matches 209; Conservative 55; Mismatches 47; Indels 17; Gaps 2;
```

```

QY 39 LEVPOEAMVVERMGRFRIIEPGLNIIIPVLDRIYVOSLKEIYINVEPOSATVLDNT 98
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MEVPOQEAMVVERMGRFRIIEPGLNIIIPVLDRIYVOSLKEIADVPKOSAITSDNT 60
QY 99 LOIDGVLYLRIMDPKASYGVDEPEYAVTOLAQTMRSELGKLSXDKYFRRESINASTV 158
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LSIDVLYLRITDPKASYGVDEPEFAITOLAQTMRSELGKMSDKYFRRESINASTV 120
QY 159 DAINOAAWCGIRCLRIRYEIKOIHVPPRYKESOMOVEARRRRAIVLESEGRESAINVA 218
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 DSINKASEPMGIACIRYERIRDRIRLPTFRHEAMOMOVEARRRRAIVLESEGREAEINVA 180
QY 219 EGKKAQOILASEAEKAEQINOAGEASAVLAKAKAEAIRILAAALTOHNGDAASITV 278
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 EGKRSRILASEAEQEHINKASGEAAAIIVADARSLALIAKSLSHLDQNASLTL 240
QY 279 AEQYVASAKLAKDSNTILLPSNPGDVTSMVAQMGVYALTKAPPGTSDLSGSSGSD 338
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 AEQYGAERKLAKTNTMLPSNPGDVNGFVAQALAVYNNHVSNS-----NOATRKSEN 293
QY 339 VOGTDA-----SXDELDRIYKMS 356
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 294 VKGVACIUMAKSVYEKLEOEDKSSVKMN 321

RESULT 8
Q9XVP9 PRELIMINARY: PRT: 334 AA.
ID Q9XVP9 01-NOV-1999 (TREMBlrel. 12, Created)
AC Q9XVP9 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F30A10.5 PROTEIN.
GN F30A10.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
   investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81072; CAB03018.1; -.
DR InterPro: IPR001107; Band 7.
DR InterPro: IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PRO0721; STOMATIN.
DR SMART; SM00244; PHB; 1.
SQ SEQUENCE 334 AA; 36704 MW; 10F98BB9A44E5ED CRC64;

Query Match 54.2%; Score 957.5; DB 5; Length 334;
Best Local Similarity 61.2%; Pred. No. 1.1e-53;
Matches 188; Conservative 58; Mismatches 46; Indels 15; Gaps 2;

QY 29 SSGLP-----RNTVLFVPOOEAMVVERMGRFRIIEPGLNIIIPVLDRIYV 76
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 SSTLPLAVTSSQAAAHNTVINEVPOOEAMVVERMGRFRIIEPGLNIIIPVLDRIYV 83
QY 77 OSLKEIYINVEPOSATVLDNTLOIDGVLYLRIMDPK---ASYGEDEPEYAVTOLAQT 133
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 QNLRELAIEIPGGAITTDNVQLRDLGVLYLRVFPDYKACDASYGVDEPEFAVTOAQT 143
QY 134 MSELGKLSXDKYFRRESINASTVDAINOAADCMGIRCLRIRYEIKDIIHVPYRKESMOM 193
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 144 MRSEVGKINLDTVFRERELNENIVFAINKASAPWGICQMRYEIRDMQPSKIOEMQMO 203

```

```

QY 194 VEAERKRATVLESEGRESAINVAEGKKAQOILASEAEQINOAGEASAVLAKAKA 253
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 VEAERKRRAIIESEGRIRAAINRAEGDKKSAIILSEAYOAEIRINAKAEAEVILAKAS 263
QY 254 KAEAIRILAAALTOHNGDAASITVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQAM 313
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 264 RAKAIERIALALEKOGANAAGLTVAEQYVGAFGNLAKSNTIVLPANLSDPGSMVSOAL 323
QY 314 GYVGALT 320
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 AVYDSL 330

RESULT 9
Q60121 PRELIMINARY: PRT: 354 AA.
ID Q60121 01-AUG-1998 (TREMBlrel. 07, Created)
AC Q60121 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 39.3 KDA PROTEIN C1665.07C IN CHROMOSOME II.
GN SPBC1665.07C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M.; Wood V.; Rajandream M.A.; Barrell B.G.; Beck A.;
RA Reinhardt R.;
RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
CC -i- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
DR EMBL; AL02354; CAI19027.1; -.
DR InterPro: IPR001107; Band 7.
DR InterPro: IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR SMART; SM00244; PHB; 1.
DR PROSITE; PS01270; BAND_7; FALSE_NEG.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39274 MW; 66B09E6A12BDC030 CRC64;

Query Match 48.8%; Score 861.5; DB 3; Length 354;
Best Local Similarity 56.7%; Pred. No. 1.7e-47;
Matches 170; Conservative 56; Mismatches 71; Indels 3; Gaps 2;

QY 35 NNTVLFVPOOEAMVVERMGRFRIIEPGLNIIIPVLDRIYVOSLKEIYINVEPOSATV 94
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 50 NNTIKVPOOVAVVERMGRFRSILTPGVAFLAPITDKIAYINSLERALEIPTQSAITL 109
QY 95 DNVTLOIDGVLYLRIMDPKASYGVDEPEYAVTOLAQTMRSELGKLSXDKYFRRESIN 154
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 DNVTSLGLDGVLYIQVDPKASYGVDEADYVAISOLAQTMRSEIGRLTDHVLREHQSIN 169
QY 155 ASITVAINOAAADCMGIRCLRIRYEIKDIIHVPYRKESMOMOVEARRRRAIVLESEGRES 214
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 IHTTDAINKAASWGRICRIRHEIRDIRPESYVMAHQOVSARORAKAEILSEGRKOA 229
QY 215 INVAEGKKAQOILASEAEKAEQINOAGEASAVLAKAKAEAIRILAAAL-TOHNGDA 273
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 230 INVAEGDKAEILDSRGQIKITINSALAEQAIRERASATASITAVLADSIKQEGLEA 289
QY 274 ASITVAEOYVSASFSLAKDSNTILLPSNPGDVTSMVAQMGVYALTK--APVPGTSD 331
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 VSLYIAQYITNGLAKASNSMIVPASTSDVSGVVAQALSIKQVSKTAPDKSTPKEL 349

RESULT 10
Q93VP9 PRELIMINARY: PRT: 411 AA.
ID Q93VP9 01-NOV-1999 (TREMBlrel. 12, Created)
AC Q93VP9 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F30A10.5 PROTEIN.
GN F30A10.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
   investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81072; CAB03018.1; -.
DR InterPro: IPR001107; Band 7.
DR InterPro: IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PRO0721; STOMATIN.
DR SMART; SM00244; PHB; 1.
SQ SEQUENCE 354 AA; 39274 MW; 66B09E6A12BDC030 CRC64;

```

AC 093VP9.
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 45.0 KDA PROTEIN.
GN ATG27580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Yu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Natsume M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinozaki K., Shinozaki A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4g27580 (GI:7269612).";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamita A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Natsume M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinozaki K., Shinozaki A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4g27580 (GI:7269612).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY0349109; AAL15215.1; -
DR EMBL; AY034924; AAK59431.1; -
KW Hypothetical protein.
SQ SEQUENCE 411 AA; 45020 MW; 94E24ED4C9889633 CRC64;

Query Match 46.7%; Score 826; DB 10; Length 411;
Best Local Similarity 50.1%; Pred. No.3.8e-45;
Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;
OY 29 SSGLEPRNTVTL-----EYPOQEAANVERMGRFHRLLEGILIPVLDRIYVOSL 79
DB 45 SAGYPSNSFOLTPTPTNMGIRIVPERKAFVIERFGKYATTLPSGIHFLIPFVDRIYVHSL 104
OY 80 KEIVINPEQSAVLTNDYNTLOIDGVLVLRIMDPYKASGVEDPEYAVQOLAQTTRSELG 139
DB 105 KEAIPINQTAITKDNVSIHIDGVLVYKIVDPKLASTGVESPIYAVVOLAOQTTMRSELG 164
OY 140 KLSXDKVRERESLNASIVDAINQADCGIRCLRYETIKDIHVPRVYESMOMQVEAERR 199
DB 165 KITLDTKTEERDITNEKIIVEAINVAKMDWGLQCLRYETRIDIMPPIGVRAAMEQAEERK 224
OY 200 KRAIVLESEGTRESAINVAEGKKAOIILASEAKAEQINQAAAGEASVILAKAKAAEAIR 259
DB 225 KRAQILSEEGERSHININADGKSSVILASEAKMDQVNRAGAEALILARQAAAKGLV 284
OY 260 ILAALITQHNDAASLVVAEOYVSAFSLAKDASNTILLPSNPGVTSVVAQMGVYCAL 319
DB 285 ILSQSLKGTGVEAASLVVAEOYITAFGNIAKEGIMILLPSGASNPASHMIAQALTMYSLSL 344
OY 320 TKAPVPGTPTDSLSSGSSRDVQGTASXDEELDRVK 354
DB 345 -----VINGRSKDHQETQALDFTDELELE 368

RESULT 11
O9LVMO PRELIMINARY; PRT; 515 AA.
AC 09T082;
DT 09T082;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 55.9 KDA PROTEIN.
GN T29A15.70 OR AT4G27580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Beyan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
RA Volckaert G., Hohnsels J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robben J., Gymnopoulos B., Volckaert G., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL0355602; CAB38270.1; -
DR EMBL; AL161571; CAB81408.1; -
DR InterPro; IPR001107; StomatL.
DR InterPro; IPR001972; StomatL.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATRN.
DR SMART; SM00244; PHB; 1.
KW Hypothetical protein.
SQ SEQUENCE 515 AA; 55923 MW; 420FE3DB5B7AFDBA CRC64;

Query Match 46.7%; Score 826; DB 10; Length 515;
Best Local Similarity 50.1%; Pred. No.5.1e-45;
Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;
OY 29 SSGLEPRNTVTL-----EYPOQEAANVERMGRFHRLLEGILIPVLDRIYVOSL 79
DB 45 SAGYPSNSFOLTPTPTNMGIRIVPERKAFVIERFGKYATTLPSGIHFLIPFVDRIYVHSL 104
OY 80 KEIVINPEQSAVLTNDYNTLOIDGVLVLRIMDPYKASGVEDPEYAVQOLAQTTRSELG 139
DB 105 KEAIPINQTAITKDNVSIHIDGVLVYKIVDPKLASTGVESPIYAVVOLAOQTTMRSELG 164
OY 140 KLSXDKVRERESLNASIVDAINQADCGIRCLRYETIKDIHVPRVYESMOMQVEAERR 199
DB 165 KITLDTKTEERDITNEKIIVEAINVAKMDWGLQCLRYETRIDIMPPIGVRAAMEQAEERK 224
OY 200 KRAIVLESEGTRESAINVAEGKKAOIILASEAKAEQINQAAAGEASVILAKAKAAEAIR 259
DB 225 KRAQILSEEGERSHININADGKSSVILASEAKMDQVNRAGAEALILARQAAAKGLV 284
OY 260 ILAALITQHNDAASLVVAEOYVSAFSLAKDASNTILLPSNPGVTSVVAQMGVYCAL 319
DB 285 ILSQSLKGTGVEAASLVVAEOYITAFGNIAKEGIMILLPSGASNPASHMIAQALTMYSLSL 344
OY 320 TKAPVPGTPTDSLSSGSSRDVQGTASXDEELDRVK 354
DB 345 -----VINGRSKDHQETQALDFTDELELE 368

RESULT 12
O9LVMO PRELIMINARY; PRT; 401 AA.
AC 09LVMO;
DT 01-OCT-2000 (TReMBLrel. 15, Created)

Db 79 VTLSTIDGVLVYKTIIDPMASISYVNNPYAITQLAOTMSEIGKLPDITFEERDTLVNA 138
Oy 157 IVDAINQAADCGICRLREIKDIHVPRVKESQMOMQVAEERRKATVLESEGTRESAIN 216
Db 139 IYSAINQAAMINNGICQMRKEIKDIOPOTIILKAMELOVAEROKRAOIISEGNQAKIN 198
Oy 217 VAEGKKAQOIIASEAEKAOIQNAGEASAVLAKAKAKAERIRIILAAALTOHNGDAASL 276
Db 199 HAEGERKAOIVLNSSEASYTDQVNRKAGEAIGLVATATANSIEIYMAAIOKTGSDAVAL 258
Oy 277 TVAEQVSAFSKLAKDSNTILLPSNPGDVTSMVAQMGYCAL 319
Db 259 KIAEQYISAFGNLAKDTNVLPTNLSEPSFVTGALTTFNQL 301

Oy 351 DRVK 354
Db 388 QRPK 391

Search completed: September 22, 2002, 18:30:03
Job time: 546 sec

RESULT 15
O9M585
ID O9M585 PRELIMINARY: PRT: 394 AA.
AC O9M585:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STOMATIN-LIKE PROTEIN.
CN STM1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RM [1]
RP SEQUENCE FROM N.A.
RA Nadimpalli R., Yalpani N., Johal G.S., Simmons C.R.;
RT "Prohibitins, Stomatins, and Plant Disease Response Genes Comprise a
RT Protein Superfamily that Controls Cell Proliferation, Ion Channel
RT Regulation, and Death."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF236372; AAF68388.1; -;
DR InterPro: IPR001107; Band_7;
DR Pfam: PF01145; Band_7; StomatIn.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB.1
SQ SEQUENCE 394 AA: 43277 MW: 52A409DCFI5C5A45 CRC64;

Query Match 45.0%; Score 794.5; DB 10; Length 394;
Best Local Similarity 47.0%; Pred. No. 3.7e-43;
Matches 171; Conservative 63; Mismatches 107; Indels 23; Gaps 3;

Oy 14 ABGLSTGFVPRSGRASSGL-----PRNTVLEFVPOQEAHVVERMGFRHRILEPGLNILL 67
Db 28 APALSRFSFRFNPDDSSMFDPPEPVNMGVSIVPEKKAIVVERFGKYLTGLSGFHLII 87
Oy 68 PVLDRIRYQSLKEIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDPEYAVT 127
Db 88 PAVDRIAVYHSLKEETIPIPHONAIKRDVNTIQIDSVIYVKIMDPYLASGVENPIYAVL 147
Oy 128 QLAOTTMRSELGLSKDXKVFRESINASTVDAINQAOCMGIRCLRYETIKDIHVPRVK 187
Db 148 QLAOTTMRSELGKITTDKTFEERDALNEKIVSAINEAATDMGLKCIETREIDINPPAGIR 207
Oy 188 ESMQOVAEERRKATVLESEGTRESAINVAEGKKAQOIIASEAEKAEQIQNAGEASAV 247
Db 208 QAMEKQAEAEERRKRAQIILESEKKAQOIILESEKKAQOIILESEKKAQOIILESEKKAQOII 267
Oy 248 LAKAKAKAEAIRILAAALTOHNGDAASLTVAEQVSAFSKLAKDSNTILLPSNPGDVT 307
Db 268 LAKSEATARGMRLVSDAMTTEGSAKASIKLABQYIEAFSNLAQKNTMTLLPGDSASPAS 327
Oy 308 MVNAQMGVY-----GALTKAP-----VPGTPDLSGSSRDVGTGDSXDEEL 350
Db 328 FVAQAKKTYEQIHSQALSKHPQIELEKESGETSPAPSSSEAKTTPPLIEADSNQTFSL 387

	Matches	356;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MLARAARGHMPFACGLSTGFMPBGRGRASSGLPRNTVTLFVFOQEAAMVFNENGRFRRIIE	60							
Db	1	MLARAARGHMGFFAAGLSTGFMPBGRGRASSGLPRNTVTLFVFOQEAAMVFNENGRFRRIIE	60							
QY	61	PLGLNTLILVLDIRIRVOSKREIVINWPBESATLTNNLTQIDGVLTYLRIMDPKASGYE	120							
Db	61	PLGLNTLILVLDIRIRVOSKREIVINWPBESATLTNNLTQIDGVLTYLRIMDPKASGYE	120							
QY	121	DPEYAVTOLAQTMRSELGLKLSXDRVFRRESLNASIVDAINQADCMGIFCLRYEIKDI	180							
Db	121	DPEYAVTOLAQTMRSELGLKLSXDRVFRRESLNASIVDAINQADCMGIFCLRYEIKRI	180							
QY	181	HVPFVKESMOMQVBAERRKRATVLESECTRRESAINVAEGKKQAOQLLASEAKRAEQINQA	240							
Db	181	HVPFVKESMOMQVBAERRKRATVLESECTRRESAINVAEGKKQAOQLLASEAKRAEQINQA	240							
QY	241	AGEASAVLAKKAKKAEAIRIILAAALTQHNHGDAASLTVAEQVYSAFSKLAKPSNTLILFS	300							
Db	241	AGEASAVLAKKAKKAEAIRIILAAALTQHNHGDAASLTVAEQVYSAFSKLAKPSNTLILFS	300							
QY	301	NPGDVTSWVAQMGVYGALTAKPVGCTPDSLSSGSSRDVGQFDASDEDELDIRYKMS	356							
Db	301	NPGDVTSWVAQMGVYGALTAKPVGCTPDSLSSGSSRDVGQFDASDEDELDIRYKMS	356							

Query Match 26.7%; Score 472; DB 1; Length 381;
 Best Local Similarity 33.2%; Pred. No. 2.9e-38;
 Matches 110; Conservative 72; Mismatches 123; Indels 26; Gaps 6;

```

OY      38 VLEVPOQEAAMVYERNGRRHRIIEPGILNIIIPYLDRIKRYQSLSKEIYINVEQSAVTLDNV 97
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      26 VALTPOAEAAVYETIGRRYSRTVSGQLTTLVPIDRAVARDLRRERVSPPPVPTIDNL 85
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      98 TLQIDGVLYLRIMDPYKASGVGEDEPEYAVTOLAQTMRSELKLSXDKYFRRESLNAST 157
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      86 TLINDTVVYEGYTVQAANVELSNITVGEQTLTTTLTRVNVGGMILEQTLTSRDQNAQL 145
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      158 VDAINQAADCMGICRCLRIEIKDIIHPPRYKESOMQOVEAEKRRKATVLESEGTRESALNV 217
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      146 RGVLEAARGMGLRAVARVELRSIDPPSTIQASMEKQMKADREKRRAMILLTSEGTREAIINQ 205
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      218 AEGKKRAQ-----ILASEAKAEQIINQAAGEASAVLAKAKAKAEAIRILLAAALT 266
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      206 AEGCKQAQIILAEAGAKQAQAIILAAEDROSRMIRAQGERAAATVIAQAGKAKALEKTPAAL- 264
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      267 QHNGGAASLTVAEYGVSAFSKLAK-DSNTI-LTSSNPEDVYSMAQAAGVYG----AL 319
        | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      265 --KAGRPPTPEMLATVYLOTLTPEMAGKDAKVVAVSDSFNAALQGTFLRLGKRGEDGVFFR 322

```

RESULT 2
 US-08-781-562-6
 Sequence 6, Application US/08781562
 Patent No. 5763589
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Goll, Surya K.
 TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Potter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTED for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,562
 FILING DATE: Herewith
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0181 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 381 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: OMI
 CLONE: 79701
 US-08-781-562-6

RESULT 3
 US-08-781-562-7
 Sequence 7, Application US/08781562
 Patent No. 5763589
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Goli, Suyra K.
 TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESS: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,562
 FILING DATE: Herewith
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0181 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ. ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 199 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:24:54 : Search time 24.86 Seconds
(without alignments)
349.779 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767
Sequence: 1 MLRAARGHMGFAEGSLTG:.....RDVGTDSXDELDYKMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	99.8	356	1	US-08-781-562-1
2	472	26.7	381	1	US-08-781-562-6
3	354.5	20.1	199	1	US-08-781-562-7
4	281.5	15.9	288	1	US-08-781-562-3
5	271	15.3	280	1	US-08-781-562-4
6	169	9.6	415	1	US-08-781-562-5
7	154	8.7	379	1	US-08-279-270A-1
8	104.5	5.9	516	2	US-08-762-106-8
9	104.5	5.9	516	3	US-08-745-404-2
10	104.5	5.9	516	4	US-09-320-774-8
11	104.5	5.9	527	2	US-08-762-106-9
12	104.5	5.9	527	4	US-09-320-774-9
13	104.5	5.9	552	3	US-08-745-404-3
14	104	5.9	1339	4	US-09-310-187A-1
15	99.5	5.6	564	4	US-08-115-746-2
16	99.5	5.6	564	4	US-09-115-746-2
17	99.5	5.6	643	2	US-08-216-894-8
18	99.5	5.6	643	4	US-09-115-746-8
19	99.5	5.6	1886	4	US-08-938-105-3
20	99	5.6	1792	2	US-08-962-284-4
21	97.5	5.5	472	2	US-08-216-894-10
22	97.5	5.5	472	4	US-09-115-746-10
23	97.5	5.5	1196	1	US-08-144-121-4
24	97.5	5.5	1196	2	US-08-735-893-4
25	97	5.5	400	3	US-08-938-830-29
26	96.5	5.5	914	4	US-09-085-199B-4
27	95.5	5.4	3170	2	US-07-642-734C-5

28	95.5	5.4	3170	3	US-08-439-009A-5	Sequence 5, Appl
29	95	5.4	1184	4	US-09-541-782-2	Sequence 2, Appl
30	93.5	5.3	656	1	US-08-371-930-23	Sequence 23, Appl
31	93.5	5.3	656	5	PCT-US94-01712-23	Sequence 23, Appl
32	93.5	5.3	676	1	US-08-371-930-24	Sequence 24, Appl
33	93.5	5.3	676	5	PCT-US94-01712-24	Sequence 24, Appl
34	93.5	5.3	694	1	US-08-339-152A-18	Sequence 18, Appl
35	93.5	5.3	694	2	US-08-007-999B-5	Sequence 5, Appl
36	93.5	5.3	694	2	US-08-689-276A-5	Sequence 5, Appl
37	93.5	5.3	695	1	US-08-371-930-27	Sequence 27, Appl
38	93.5	5.3	695	1	US-08-123-702-2	Sequence 2, Appl
39	93.5	5.3	695	1	US-08-339-152A-30	Sequence 30, Appl
40	93.5	5.3	695	2	US-08-104-165-1	Sequence 1, Appl
41	93.5	5.3	695	3	US-08-464-250-1	Sequence 1, Appl
42	93.5	5.3	695	4	US-08-464-250-1	Sequence 1, Appl
43	93.5	5.3	695	4	US-09-458-481B-7	Sequence 7, Appl
44	93.5	5.3	695	4	US-09-458-481B-8	Sequence 8, Appl
45	93.5	5.3	695	5	PCT-US94-01712-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-781-562-1
; Sequence 1, Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,562
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-781-562-1
Query Match 99.8%; Score 1763; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 2, 6e-165;


```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152370.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154179.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159588.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 46.7%; Score 826; DB 21; Length 411;
Best Local Similarity 50.1%; Pred. No. 4,5e-64;
Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;

```
OY 29 SSGLPNTVVL-----FVPOQEAWVERGRHRIIEPDLNLIPIVLDRIIRVQSL 79
DB 45 saqyspsnfqlprrptmgwrlrlyperkafylerfgkyatcltspgihlrlpvdrlayvhsI 104
OY 80 KEIVIVPQSAVTLDDNVTLOIDGVLYLRIMDPYKASGVGEVDEPEVAVTOLAQTMMSELG 139
DB 105 keeaipmpqtaiktksvnhidgvlrvkltvdpkrlasgyvesiyavvglaqtlmtselg 164
OY 140 KLSXKQVFPRESLNASIVDAINQAADQWGRICRLREIKDIHVPVRYKSMQNOVEAERR 199
DB 165 kltdlktfeerdlinekivealnvaaakdwglgclryelrldimpbhgyraamemqaeatrk 224
OY 200 KRATVESEGTRESATINVAEGKQAOIILASEAKAOINQAGASAVLAKAKAKAEATR 259
DB 225 kraqllesegqrshnlndgkssvllaseaakmdqvmragaeaaallaratagatakgyIv 284
```

```
OY 260 ILAALTOHNGDAASLTVAEOVSAFSKLAKDSNTILLPSNCGDVTSMVAQMGVYGL 319
DB 285 llsqslketgyveaaslrvaeyitafgnlaketclmlpsgaspsnmlaqeltnyksi 344
OY 320 TRKAVPGTPDSLSSGSSRDVQGTDAKXDELDKRVK 354
DB 345 -----vlnpskdhgetqaldecldleele 368

RESULT 14
AAG31395
ID AAG31395 standard; Protein: 515 AA.
XX
AC AAG31395;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37695.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-01231825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130479.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136592.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
```

XX AAG50184;
AC 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63567.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
PN EPI03405-A2.
XX
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145228.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.7%; Score 826; DB 21; Length 411;
Best Local Similarity 50.1%; Pred. No. 4,5e-64;
Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;

QY 29 SSGLPRTNVL-----FVPOEAMVVERMGFRHRIIEPGTILIPYLDRIIRYQSL 79
DB 45 saypsnsfqltpptmngitlveperkafterfgkyatlpsghiflptvdrlyahsl 104
QY 80 KEIVINPEQSAVTLNDNTLQIDGVLYLRIMDPYKASYGVDEPEYAVTOLAQTTRSELG 139
DB 105 keeipipnqfatlkdvshidgvlyvkiwdpklaesyvespiyavvqtaqtumzslg 164
QY 140 KLSMDKFRERESIMASIVAINQAMADCMGIRCLRIRYIKIKIHVPRKESMOMQVEAEER 199
DB 165 kitldkltfeerdtnekiivaavaakdwlgiclyelrldimphyvraamengaeerlk 224
QY 200 KRATVLESEGTRESAINVAGCKKQAOITLASEAEKAEQINQAGRASAVLAKAKAKAEAIR 259
DB 225 kragllesegerghindagkksvllaeseaakmdgvnraagaeallaraqetavg 284
QY 260 ILAAILTOHNGDAASLTVAEQYVSAPSKLAKDSNTLLPSNPGDVTSMVAQMGVYGAL 319
DB 285 llsqllketgveaasirvaeqyitafgnlakegtlmlpsgaenpsamiateltmksl 344
QY 320 TKAPVPGTPOSLSSGSRDQGTDAKXDEELDRVK 354
DB 345 -----vlnpskdhqetqaldetdele 368

RESULT 13
AAG50184
ID AAG50184 standard; Protein; 411 AA.

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of hematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.

XX Sequence 463 AA:

Query Match 53.5%; Score 946; DB 22; Length 463;
Best Local Similarity 64.0%; Pred. No. 1,6e-74;
Matches 242; Conservative 20; Mismatches 74; Indels 42; Gaps 19;

OY 13 FAEG--LSTGFMRSGRASSGLPRNMYLVLPPOEAMVVERMRPHRIIEGNIILPVL 70
DB 94 fLggsllaaagrapr--rassgdlprncvllfpqgeawvermrtrfhllepjlnllpvl 151
OY 71 DRIRYV-QSLKETIV-INVP-EGSAYLDNVYLIQIDGVLYLRINDPYKAS-YGEDEPEYAV 126
DB 152 drrlyraasskeiclsfnvalnsavcldnvclqldgvlylrindpykqkrlgyedpeyav 211
OY 127 TQLAQT-MRSELGKLSXKXVFERESLNASTYDAINQAAOCIGICRLREIDHWPPR 165
DB 212 lqpslnhvdqsglslldlfrereslnaslydaIngaacwglrlcrlryelldhnprr 271
OY 186 VKESMOMOVEAERR-----KRAVLESERG--TRESAINVAGCKKQAOVLASEAEKAEQ 236
DB 272 vke-----vxagmhggggqgqgktratlvesegdpkksaqlwgeqkqgqlaseaeaeq 326
OY 237 INQAGEASAVLAKAR-----AKAEAIRLMAALTQHNGDAASLT--VAE--QYVSASF 287
DB 327 Insgqerascsgeprkllkalftrfgaldpknwelaafxlvaeaamfsgvfp 386
OY 288 KLA--KDSMTI--LLEPSNGDVTSMAQAMGV-YGALTAKAPVPGTDP---SLSSGSRDY 339
DB 387 klqprnsgqlldclalpt-pgdvltsmvaagpwwxygaltspspsardfqlhssgssrdv 445
OY 340 OGTDAS-XDELDPRVKMS 356
DB 446 ggtldaspxxgnllrvkms 463

RESULT 12

AAC20661
ID AAC20661 standard: Protein: 411 AA.

AC AAG20661;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 22939.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 9905-0121825.

PR

PR 05-MAR-1999; 9905-0123180.

PR

PR 09-MAR-1999; 9905-0123548.

PR

PR 23-MAR-1999; 9905-0125788.

PR

PR 25-MAR-1999; 9905-0126264.

PR

PR 29-MAR-1999; 9905-0126785.

PR

PR 01-APR-1999; 9905-0127462.

PR

PR 06-APR-1999; 9905-0128234.

PR

PR 16-APR-1999; 9905-0129845.

PR

PR 19-APR-1999; 9905-0130077.

PR

PR 21-APR-1999; 9905-0130449.

PR

PR 23-APR-1999; 9905-0130510.

PR

PR 23-APR-1999; 9905-0130891.

PR

PR 28-APR-1999; 9905-0131449.

PR

PR 30-APR-1999; 9905-0132048.

PR

PR 30-APR-1999; 9905-0132407.

PR

PR 04-MAY-1999; 9905-0132484.

PR

PR 05-MAY-1999; 9905-0132485.

PR

PR 06-MAY-1999; 9905-0132486.

PR

PR 07-MAY-1999; 9905-0132487.

PR

PR 11-MAY-1999; 9905-0134256.

PR

PR 14-MAY-1999; 9905-0134218.

PR

PR 14-MAY-1999; 9905-0134219.

PR

PR 14-MAY-1999; 9905-0134221.

PR

PR 18-MAY-1999; 9905-0134370.

PR

PR 19-MAY-1999; 9905-0134768.

PR

PR 20-MAY-1999; 9905-0134841.

PR

PR 21-MAY-1999; 9905-0135124.

PR

PR 21-MAY-1999; 9905-0135353.

PR

PR 24-MAY-1999; 9905-0135629.

PR

PR 25-MAY-1999; 9905-0136021.

PR

PR 27-MAY-1999; 9905-0136392.

PR

PR 28-MAY-1999; 9905-0136782.

PR

PR 01-JUN-1999; 9905-0137222.

PR

PR 03-JUN-1999; 9905-0137528.

PR

PR 04-JUN-1999; 9905-0137502.

PR

PR 07-JUN-1999; 9905-0137724.

PR

PR 08-JUN-1999; 9905-0138094.

PR

PR 10-JUN-1999; 9905-0138540.

PR

PR 10-JUN-1999; 9905-0138847.

PR

PR 14-JUN-1999; 9905-0139119.

PR

PR 16-JUN-1999; 9905-0139452.

PR

PR 16-JUN-1999; 9905-0139453.

PR

PR 17-JUN-1999; 9905-0139452.

PR

PR 18-JUN-1999; 9905-0139454.

PR

PR 18-JUN-1999; 9905-0139455.

PR

PR 18-JUN-1999; 9905-0139456.

PR

PR 18-JUN-1999; 9905-0139457.

PR

PR 18-JUN-1999; 9905-0139458.

PR

PR 18-JUN-1999; 9905-0139459.

PR

PR 18-JUN-1999; 9905-0139460.

PR

PR 18-JUN-1999; 9905-0139461.

PR

PR 18-JUN-1999; 9905-0139462.

PR

PR 18-JUN-1999; 9905-0139463.

PR

PR 18-JUN-1999; 9905-0139750.

PR

PR 18-JUN-1999; 9905-0139750.

PR

PR 21-JUN-1999; 9905-0139763.

PR

PR 22-JUN-1999; 9905-0139817.

PR

PR 22-JUN-1999; 9905-0139899.

PR

PR 23-JUN-1999; 9905-0140353.

PR

PR 23-JUN-1999; 9905-0140354.

PR

PR 24-JUN-1999; 9905-0140695.

PR

PR 28-JUN-1999; 9905-0140891.

PR

PR 29-JUN-1999; 9905-0140991.

PR

PR 30-JUN-1999; 9905-0141287.

PR

PR 01-JUL-1999; 9905-0141842.

OY 339 VQGTDA-----SXDEELDRVKMS 356
 DB 294 VKYVGACINAKSVKEYKEJGQKSSVKMN 321

RESULT 10
 AAU33158
 ID AAU33158 standard; Protein: 222 AA.

AC AAU33158:

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3649.

XX Human: vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic

PS vaccination, testing and therapy -

PS Claim 20; Page 716; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 222 AA:

Query Match 57.2%; Score 1011; DB 22; Length 222;

Best local similarity 92.9%; Pred. No. 1, 1e-80; Indels 4; Gaps 2;

Matches 208; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

OY 1 MLARAARAHMGPFAEG--LSTGFMPRSGRASGLPRTNVTLVFVPOOEAVVBERMGRFHR1 58
 DB 1 MLARAARAHMGPFAEG--LSTGFMPRSGRASGLPRTNVTLVFVPOOEAVVBERMGRFHR1 58
 OY 59 LEPGLNLLPVDLRIRYVQSLKELIVNPPQSAVTLDNVTLDQVLYLRIMDPKASYG 118

DB 59 LEPGLNLLPVDLRIRYVQSLKELIVNPPQSAVTLDNVTLDQVLYLRIMDPKASYG 118
 OY 119 VEDPEYAVTOLAQTMRSELKLSXDKVFERERESINASTVAINOADQWGRICRLRYEIK 178
 DB 119 VEDPEYAVTOLAQTMRSELKLSXDKVFERERESINASTVAINOADQWGRICRLRYEIK 178
 OY 179 DIHPPRYKESHQMOQVEARRKRAIVLESEGTRESAINVAEKK 222
 DB 179 dihprrvkesmqmqvgaerkracvlesegtresainvaeqk 222

RESULT 11

AAU28241
 ID AAU28241 standard; Protein: 463 AA.

AC AAU28241:

DT 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 598.

XX Human: secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

OS Homo sapiens.

PN WO200166689-A2.

PD 13-SEP-2001.

PF 05-MAR-2001; 2001WO-US04942.

PR 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0618447.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

DR WPI: 2001-589934/66.

DR N-PSDB: AAS45141.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -

XX Example 2; SEQ ID No 598; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

OY 299 PSNPGDVTSMVAQAMGYGALTKAPVPGTSPDLSGSSSRDVGOTDASXDEELDRVKMS 356
|||||
Db 299 psnpgdvtsmwagangyvgaltkapyvpgtpdlsissgsstidvgtdasldeeldrvkmt 356

RESULT 7
AAM78566
ID AAM78566 standard; Protein: 358 AA.

XX AAM78566;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1228.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX MO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0663561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;

XX WPI: 2001-476283/51.

XX N-PSDB: AAK51699.

XX Nucleic acids encoding polypeptides with cytokine-like activities;
PT useful in diagnosis and gene therapy -

XX Claim 20; Page 3486-3487; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK51435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 358 AA;

Query Match 92.5%; Score 1635; DB 22; Length 358;
Best Local Similarity 94.7%; Pred. No. 4,1e-135;
Matches 341; Conservative 2; Mismatches 11; Indels 6; Gaps 3;

OY 1 MLARAAGHMGPFACG--LSTGFMFRSGRASSGLPRTNTVLFVPOQEAAMVVERMGFRHRI 58
|||||
Db 1 mlaraargtgalllrsgllaasgrapr--raasgiprntlvlfvpgqeaawvermgfrfiri 58

OY 59 LEPGLNLIIPVLDRIYVQSLKEIVINVEQSAVTLNDVTLQIDGVLYRLIMDPYKASG. 118
|||||
Db 59 lpggnllipvldriyrqslkeivinvpeqsavldnvtlqldgvlyrlimdpkasyg 118

OY 119 VEDPEYATQLAQTMRSELKLSXDKVFRFRESINASTVAINQAACWGRICLRFEIK 178
|||||
Db 119 vedpeyavtqlaqtmrseqlklsidkvfrereslnasivdaingaaacwgrlclryelk 178

OY 179 DIHVPRVKESMOMQVEARRKATVLESEGTRESAINVAEGKKAQOLTAESAERAEQIN 238
|||||
Db 179 dihvprvkessmqvvearrkratvlesegtrresainvaegkkqaqltaesaekaeqin 238

OY 239 QAAGFASAVLAKAKAKAEIRILAAALRQHNGDAASLTVAEQVYSAPSKLAKDSNTILL 298
|||||
Db 239 qaagfasavlakakakaairllaaaltqhngdaaasltvaeqysatsklakdsntlll 298

OY 299 PSNPGDVTSMVAQAMGYGALTKA--PYGTPDLSGSSSRDVGOTDASXDEELDRVKMS 356
|||||
Db 299 psnpgdvtsmwagangyvgaltkapyvpgtpdlsissgsstidvgtdasldeeldrvkms 356

RESULT 8

AAW78284
ID AAM78284 standard; Protein: 306 AA.

XX AAM78284;

XX 13-APR-1999 (first entry)

XX Fragment of human secreted protein encoded by gene 35.

XX Human: secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9856804-A1.

XX 17-DEC-1998.

XX 11-JUN-1998; 98WO-US12125.

XX 02-OCT-1997; 97US-0061060.

XX 13-JUN-1997; 97US-0049547.

XX 13-JUN-1997; 97US-0049548.

XX 13-JUN-1997; 97US-0049549.

XX 13-JUN-1997; 97US-0049550.

XX 13-JUN-1997; 97US-0049606.

XX 13-JUN-1997; 97US-0049607.

XX 13-JUN-1997; 97US-0049608.

XX 13-JUN-1997; 97US-0049609.

XX 13-JUN-1997; 97US-0049610.

XX 13-JUN-1997; 97US-0050566.

XX 13-JUN-1997; 97US-0050901.

XX 08-JUL-1997; 97US-0052989.

XX 18-AUG-1997; 97US-0051919.

XX 12-SEP-1997; 97US-0058665.

XX 12-SEP-1997; 97US-0058666.

XX 12-SEP-1997; 97US-0058668.

XX 12-SEP-1997; 97US-0058670.

PD 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-457740/49.
DR N-PSDB: ABA09225.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
XX Claim 20; Page 291; 1963pp; English.
PS
PS Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, and hence
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
XX Sequence 378 AA:

Query Match 93.2%; Score 1646; DB 22; Length 378;
Best Local Similarity 95.3%; Pred. No. 4,8e-136;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

OY 1 MABARAGHMGPPAEG--ISTGFWPSGRASSGLPNTVWLEFVPOOEAWVEMGRFHHI 58
DB 23 maraarggalllrgsllassgrapr--rassglprncvllrvpggaavemmgrfhhrl 80
OY 59 LEFGAILLPVLDRIRYVSLKEIVINVEPOSAVTLNDNTLQIDGLYLIRIMDPYASYG 118
DB 81 lepglllllpvldrlryvsglkeivinvpegasvtdlnvrlgldgylrlrmdpyasyg 140

OY 119 VEDPEYAVTQIAQTTMRSELGKLSXDKYFRERESINASIYDAINAAADQWGRICLRYEIK 178
DB 141 vedpeyavtqlaqtlmrseqlgklsldkyfreresinaslyvdaingadcwgriclrlyeik 200
OY 179 DIHVPPRVKESQKQOVVEERRKATVLESEGRRESAINVAEEKKQAOIILASAEKAEQIN 238
DB 201 dihvpprvkesmqvveerrkatvlesegrrresainvaeqkkqagqllasaeeakeeqin 260
OY 239 QAAGASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSFESKLAKDSNTJLL 298
DB 261 qaagaasavllakakaeairilaaal tqhngdaaasltvaeqysafesklakdsntjll 320
OY 299 PSNPGDVTSMVAQAMGVGALTAKAPVPGTIPDLSGSSSDVVGCTDASXDEILDYKMS 356
DB 321 psnpgdvtismvagangvygaltkapvpgtpdlssgssrdvvgctdasideildrykms 378
RESULT 5
AAM79550 standard; Protein; 378 AA.
ID AAM79550;
AC AAM79550;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 3196.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200157190-A2.
PN
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX N-PSDB: AAK52683.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
XX Claim 20; Page 281-282; 6221pp; English.
PS
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX Goli SK, Hillman JL;
 XX MPI: 1998-347418/30.
 DR N-PSDB: AAV28867.
 XX
 XX DNA encoding human integral membrane protein - useful for producing
 PT recombinant protein, for treatment of anaemia and cancer
 PS
 XX Claim 1: Column 35-38; 33pp; English.
 XX
 CC The present sequence represents human integral membrane protein (IMP).
 CC IMP may be administered to a subject to treat disorders associated
 CC with abnormal ion transport or membrane conductance as well as a
 CC variety of tumours, e.g. haemolytic anaemias and prostate, breast and
 CC pancreatic tumours. A vector capable of expressing IMP, or a fragment
 CC or a derivative thereof, may also be administered to a subject to treat
 CC the haemolytic anaemias and prostate, breast and pancreatic tumours.
 CC
 SQ Sequence 356 AA:

Query Match 99.8%; Score 1763; DB 19; Length 356;
 Best Local Similarity 100.0%; Pred. No. 2,4e-146;
 Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLARAAGHMGPFPAEG-ISTGFMPRSGRASSGLPRNTVVLFPQOEAAMVERHGRFHRI 60
 DB 1 mlaraaghwgpfpaeg-istgfmpgrassglprntvvlfpqgeawverngfrfhrlle 60
 QY 61 PGLNIIIPVLDRIYVQSLEKEIVINVEOSAVTLDNVTLDIGVLYLRIMDPYKASYGE 120
 DB 61 pglniilipvldriyvslekeivinvpegsavtldnvtldigvlylrmdpykasyge 120
 QY 121 DEEYANTOLAOITMRESEGLSKDXKFERRESINASTIVAINOACWGRICLRVEIKKI 180
 DB 121 deeyantqlaqltmrseglskdxkferresinasivainoacwgriclrveikki 180
 QY 181 HPPPRKESMOMQOAEERRKRAVLESECTRESAINVAEGKKOQTLASEAERKAEIOQA 240
 DB 181 hppprkesmqmovaerrkravlesectresainvaeqkkqagllaeseakeaginqa 240
 QY 241 AGEASAVLAKAKAKAPATITLAAITLQHNDAASLTVAEQVYSASFSLAKISNTILPS 300
 DB 241 ageasavlakakakeaalilaalqhnгдаaasltvaeqysatsklakdsntilips 300
 QY 301 NQGDVSMVAQAMGVYGAITKAPVPGTSPSLSSGSRDVGOTDASXDELDHYKMS 356
 DB 301 npgdvsmvaqamgygaaltkapvpctpslssgsrtdvgotdasxdeeldivkms 356

RESULT 2
 ID AAW78160 standard; Protein: 356 AA.
 XX AAW78160;

DT 13-APR-1999 (first entry)

DE Human secreted protein encoded by gene 35 clone HTXCS21.

XX Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 XX MO9856804-AA.
 PN

XX 17-DEC-1998.
 PD 98WO-US12125.
 XX 11-JUN-1998;
 PF 97US-0061060.
 XX 02-OCT-1997;
 PR 97US-0049547.
 PR 13-JUN-1997;
 PR 97US-0049548.
 PR 13-JUN-1997;
 PR 97US-0049549.
 PR 13-JUN-1997;
 PR 97US-0049550.
 PR 13-JUN-1997;
 PR 97US-0049606.
 PR 13-JUN-1997;
 PR 97US-0049607.
 PR 13-JUN-1997;
 PR 97US-0049608.
 PR 13-JUN-1997;
 PR 97US-0049609.
 PR 13-JUN-1997;
 PR 97US-0049610.
 PR 13-JUN-1997;
 PR 97US-0049611.
 PR 13-JUN-1997;
 PR 97US-0050566.
 PR 13-JUN-1997;
 PR 97US-0050901.
 PR 13-JUN-1997;
 PR 97US-0052989.
 PR 08-JUL-1997;
 PR 97US-0051919.
 PR 18-AUG-1997;
 PR 97US-0055984.
 PR 12-SEP-1997;
 PR 97US-0058665.
 PR 12-SEP-1997;
 PR 97US-0058668.
 PR 12-SEP-1997;
 PR 97US-0058669.
 PR 12-SEP-1997;
 PR 97US-0058750.
 PR 12-SEP-1997;
 PR 97US-0058971.
 PR 12-SEP-1997;
 PR 97US-0058972.
 PR 12-SEP-1997;
 PR 97US-0058975.
 PR 02-OCT-1997;
 PR 97US-0060834.
 PR 02-OCT-1997;
 PR 97US-0060844.
 PR 02-OCT-1997;
 PR 97US-0060855.
 PR 02-OCT-1997;
 PR 97US-0061059.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Ehnert R, Ferrite AM, Feng P, Greene JM, Lafleur DM;
 PI Moore PA, NI J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;

DR MPI: 1999-080881/07.
 DR N-PSDB: AAX04345.

PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders

PS Claim 11: Page 282-283; 380pp; English.

XX This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAX04302) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).

XX Sequence 356 AA:

QY 1 MLARAAGHMGPFPAEG-ISTGFMPRSGRASSGLPRNTVVLFPQOEAAMVERHGRFHRI 58
 Query Match 93.2%; Score 1646; DB 20; Length 356;
 Best Local Similarity 93.3%; Pred. No. 4,4e-136;
 Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:22:14 ; Search time 61.48 Seconds
(without alignments)
643.173 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767

Sequence: 1 MLARAARGHMGFPAEGLSTG.....RDVQSTASXDELDLRVMS 356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	99.8	356	19	AAW57232
2	1646	93.2	356	20	AAW78160
3	1646	93.2	356	22	AAW92559
4	1646	93.2	378	22	ABBI1981
5	1646	93.2	378	22	AAW79550
6	1643	92.5	3007	22	AAU28053
7	1635	92.5	358	22	AAW78566
8	1493	84.5	306	20	AAW78284
9	1064.5	60.2	323	22	ABW59541
10	1011	57.2	222	22	AAU33158
11	946	53.5	463	22	AAU28241

12	826	46.7	411	21	AAW20661	Arabidopsis thalia
13	826	46.7	411	21	AAW50184	Arabidopsis thalia
14	826	46.7	515	21	AAW31395	Arabidopsis thalia
15	816.5	46.2	382	21	AAW48817	Arabidopsis thalia
16	816.5	46.2	401	21	AAW48816	Arabidopsis thalia
17	816.5	46.2	411	21	AAW48815	Arabidopsis thalia
18	794.5	45.0	394	21	AAW84663	Amino acid sequenc
19	716	40.5	315	21	AAW58555	N. meningitidis am
20	716	40.5	315	21	AAW58558	N. meningitidis am
21	716	40.5	315	21	AAW25625	N. meningitidis am
22	716	40.5	315	21	AAW25628	N. meningitidis am
23	716	40.5	315	21	AAW74989	Neisseria meningit
24	716	40.5	315	21	AAW74992	Neisseria meningit
25	716	40.5	315	21	AAW75780	Neisseria meningit
26	716	40.5	315	21	AAW75783	Neisseria meningit
27	716	40.5	315	21	AAW75886	Neisseria strain z
28	716	40.5	315	21	AAW75900	Neisseria strain z
29	715.5	40.5	315	21	AAW75898	Neisseria strain z
30	714.5	40.4	315	21	AAW75889	Neisseria strain z
31	714.5	40.4	315	21	AAW75889	Neisseria strain z
32	714.5	40.4	315	21	AAW75890	Neisseria strain z
33	714.5	40.4	315	21	AAW75891	Neisseria strain z
34	714.5	40.4	315	21	AAW75901	Neisseria strain z
35	712.5	40.3	315	21	AAW58536	N. meningitidis am
36	712.5	40.3	315	21	AAW58526	N. meningitidis am
37	712.5	40.3	315	21	AAW74991	Neisseria meningit
38	712.5	40.3	315	21	AAW75781	Neisseria meningit
39	712.5	40.3	315	21	AAW75887	Neisseria strain z
40	712.5	40.3	315	21	AAW75893	Neisseria strain z
41	712.5	40.3	315	21	AAW75894	Neisseria strain z
42	712.5	40.3	315	21	AAW75895	Neisseria strain z
43	712.5	40.3	315	21	AAW75896	Neisseria strain z
44	712.5	40.3	315	21	AAW75897	Neisseria strain z
45	712.5	40.3	315	21	AAW75899	Neisseria strain z

ALIGNMENTS

RESULT 1						
ID	AAW57232	standard:	Protein:	356	AA.	
AC	AAW57232:					
XX						
DT	03-AUG-1998	(first entry)				
XX						
DE	Human integral membrane protein.					
XX						
KW	Human; integral membrane protein; IMP; cancer; anaemia; prostate;					
KW	breast; pancreatic; tumour; ion transport.					
XX						
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
FT	Misc-difference 143	/label="unknown				
FT		/note="encoded by NTC"				
FT	Misc-difference 173	/note="encoded by CTN"				
FT		/note="encoded by NTC"				
FT	Misc-difference 346	/label="unknown				
FT		/note="encoded by NTC"				
XX						
PN	US5763589-A.					
PN						
PD	09-JUN-1998.					
XX						
PF	09-JAN-1997.	97US-0781562.				
XX						
PR	09-JAN-1997.	97US-0781562.				
XX						
PA	(INCY-)	INCYTE PHARM INC.				

Search completed: September 22, 2002, 18:31:39
Job time: 532 sec

XX 11-JUN-1998; 98WO-US12125.
XX 02-OCT-1997; 97US-0061060.
XX 13-JUN-1997; 97US-0049547.
PR 13-JUN-1997; 97US-0049548.
PR 13-JUN-1997; 97US-0049549.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049606.
PR 13-JUN-1997; 97US-0049607.
PR 13-JUN-1997; 97US-0049608.
PR 13-JUN-1997; 97US-0049609.
PR 13-JUN-1997; 97US-0049610.
PR 13-JUN-1997; 97US-0049611.
PR 13-JUN-1997; 97US-0050566.
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0053984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058668.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060865.
PR 02-OCT-1997; 97US-0061059.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX
XX
DR WPI: 1999-080881/07.
DR N-PSDB: AAX04345.
XX
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Disclosure: Page 35; 380pp; English.
XX
XX
CC This sequence represents a fragment of a secreted human protein encoded
CC by the nucleic acid molecule detailed in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a
CC human immunoglobulin Fc portion (e.g. AAX04302) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic
CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 86
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX04311 for described uses).
XX
SQ Sequence 26 AA:

Query Match 7.3%; Score 26; DB 20; Length 26;
Best Local Similarity 100.0%; Pred. No. 6.3e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 276 LTVAEQYVSFAFKADSNITLPSN 301
DB 1 ltvaeqyvsafsklaksdnitlpsn 26

CC human immunoglobulin Fc portion (e.g. AAX04302) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).

XX
 XX Sequence 27 AA:
 SQ

Query Match 7.6%; Score 27; DB 20; Length 27;
 Best Local Similarity 100.0%; Pred. NO. 6.9e-19;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 MOMQVEARRKATVLESEGTRESAIN 216
 ||||||||||||||||||||||||||||
 Db 1 mnmqveaerkratvlesegtresain 27

RESULT 14
 ID AAW78285 standard; Protein: 26 AA.
 AC AAW78285;
 XX
 XX 13-APR-1999 (first entry)
 DE Fragment of human secreted protein encoded by gene 35.
 XX

Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KM developmental abnormality; foetal deficiency; blood; allergy; renal;
 KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KM inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 KM
 XX Homo sapiens.
 OS
 XX
 PN WO9856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 XX 11-JUN-1998; 98WO-US12125.
 PF
 XX
 XX 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.

PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, NI J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 DR WPI: 1999-080881/07.
 DR N-PSDB: AAX04345.
 XX
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PS Disclosure; Page 35; 380pp; English.
 XX
 XX This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a
 CC human immunoglobulin Fc portion (e.g. AAX04302) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).
 CC
 XX Sequence 26 AA:
 SQ

Query Match 7.3%; Score 26; DB 20; Length 26;
 Best Local Similarity 100.0%; Pred. NO. 6.3e-18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 ASYGVDEPEYAVTQLAQTMRSELGK 140
 ||||||||||||||||||||||||
 Db 1 asygvdepeyavtqqlatmrse1gk 26

RESULT 15
 ID AAW78287 standard; Protein: 26 AA.
 AC AAW78287;
 XX
 XX 13-APR-1999 (first entry)
 DE Fragment of human secreted protein encoded by gene 35.
 XX

Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KM developmental abnormality; foetal deficiency; blood; allergy; renal;
 KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KM inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 KM
 XX Homo sapiens.
 OS
 XX
 PN WO9856804-A1.
 XX
 PD 17-DEC-1998.

XX 13-SEP-2001.
 PD 05-MAR-2001; 2001MO-US04942.
 PF 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX Tang YF, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao Qa, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 DR WPI: 2001-589934/66.
 N-PSDB: AAS45141.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 PS Example 2: SEQ ID No 598: 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 XX
 S0 Sequence 463 AA:

Query Match 14.0%; Score 50; DB 22; Length 463;
 Best Local Similarity 100.0%; Pred. No. 2.7e-40; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0;

QY 27 RASSGLPRTNTVLEVPQGEAWVERBMRGRFRILEPGLNIPVLDIRRYV 76
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 108 rassgiprntvlfvpgqgawvercmgrfrhrllepgllnlpvlidirryv 157

RESULT 13

AAW78286
 ID AAW78286 standard: Protein; 27 AA.
 XX
 AC AAW78286;
 XX
 DT 13-APR-1999 (first entry)
 DE
 DE Fragment of human secreted protein encoded by gene 35.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; osteoclast; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN W09856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 PE 11-JUN-1998; 98MO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HDMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Edner R, Pettie AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 DR WPI: 1999-080881/07.
 DR N-PSDB: AAX04345.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Disclosure: Page 35; 380pp; English.
 CC
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a


```

XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US08656.
XX PR 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy -
XX PS Claim 20; Page 716; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
XX CC polypeptides and antibodies to the polypeptides are useful for
XX CC determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or
XX CC physiological interactions of the polypeptide. Vectors comprising
XX CC the nucleic acids encoding the polypeptides and cells genetically
XX CC engineered to express them are also useful for producing the proteins.
XX CC The proteins are useful in genetic vaccination, testing and
XX CC therapy, and can be used as nutritional supplements. They may be used to
XX CC increase stem cell proliferation; to regulate haematopoiesis; and in
XX CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX CC immune suppression and/or stimulation; as anti-inflammatory agents; and
XX CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
XX CC sequences of novel human secreted proteins of the invention.
XX SQ Sequence 222 AA;

Query Match 32.6%; Score 116; DB 22; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e-105;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 RASSGLPRNTVVLVFPQCEAMVVERMGRPHLEPGLNLIIPVLDRIRYVQSLKEIVINV 86
DB 27 rassglprntvvlfpqgeawvermgrfhrilpghllipvlidrliryvqslkeivlnv 86
OY 87 PEQSAVTLDNNTLQIDGVLXLRIMDPKASGVDEDEPXAANVQLAQTMRSELGKIS 142
DB 87 pegsavltdnntlvldgvlxlrimdpkasygvedepayavqlaqlmrselgkis 142

RESULT 11
ID AA009538 standard; Protein: 92 AA.
XX AA009538;
XX AC AAC09538;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 23430.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.

```

```

XX OS WO200164835-A2.
XX PN 07-SEP-2001.
XX PD 26-FEB-2001; 2001WO-US04927.
XX PF 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-514838/56.
XX DR N-PSDB; AA189469.
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX PS Claim 20; SEQ ID NO 23430; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AA179941-AA193641) and
XX CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 92 AA;

Query Match 14.6%; Score 52; DB 22; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.9e-43;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 NTILPSPGDVTSWVAQAMGVYGLTKAPVPGPDPDSSSGSSRDVQGNDA 345
DB 30 ntllpspgdvtswvaqamgyygaltkapvpdpdpdsssgssrdvqgdas 81

RESULT 12
ID AAU28241 standard; Protein: 463 AA.
XX AAU28241;
XX AC AAU28241;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secretory protein, Seq ID NO 598.
XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
XX KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
XX KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
XX KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
XX KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
XX KW gut protection; lung; liver fibrosis; immune deficiency; infection;
XX KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
XX KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
XX KW fertility; analgesic; pain; antigen.
XX OS Homo sapiens.
XX PN WO200166689-A2.

```

PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejherman T, Goodrich R;
XX WPI: 2001-476283/51.
DR N-PSDB: AAK51699.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PS Claim 20; Page 3486-3487; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation and which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 358 AA:

Query Match 55.6%; Score 198; DB 22; Length 358;
Best Local Similarity 99.7%; Pred. No. 2,5e-185;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 RASSGLPRTVTVLFVPOQAWYVERKGRFRHLEPGLNILEPDLRIRYQSLKEIVNV 86
DB |||||||
27 rassglprtvTVLFVpqgaawvvermgtrfhrllpeglNlllpvldrlryvgsllkeivnv 86

QY 87 PQGSNVTLDNVTLQIDGVLYLRIMDPYKASYVEDPEVAVTOLAQTMMSELGKLSXDKV 146
DB |||||||
87 pegasvnlDNvltldgVlylrImdpYkasyVEDpeVavtqlaqtmmseLgklsIdkv 146

QY 147 FRERESLNASTYDAINQOADCWCIGRCLREIKDIHVPVRKFSMOMQVFAERKRATVE 206
DB |||||||
147 ftereslnastYdaINqoaDcWcigrclREIKdiHvpvrKfsmomQvfaERkrATve 206

QY 207 SEGTRESAIVNAEGKKQAQIILASEAEKAEQINQAGASAVLAKAKAKAEAIRIILAAALT 266
DB |||||||
207 segtresaivNAegKKQAqiILaseAEkAEqINqAGasAVlAkakAKAEaiRIILAAalt 266

QY 267 OHNGGAAASTYVAEOYVSFSKLAKDSNTILPSPNGDVTSMVAQAMGYGALTAPVP 325
DB |||||||
267 ohngGAAASTyVAEOyVSfSKlAKdsNTilPspngDvtSMvAQaMGyGaltAPvp 325

QY 267 qingdaaasltvaegysfsklakdsntilpsnpgdvtsmvagamygaltkapvp 325
DB |||||||

RESULT 9
AAU33159
ID AAU33159 standard; Protein: 2797 AA.
XX
XX AAU33159;
AC
DT 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #3650.
DE
XX
XX Human: vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW Immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX

OS Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001WO-US08656.
XX
XX 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-611725/70.
DR
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
PT
XX
XX Claim 20; Page 716; 765pp; English.
PS
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC Immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 2797 AA:

Query Match 34.6%; Score 123; DB 22; Length 2797;
Best Local Similarity 100.0%; Pred. No. 4,5e-111;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 QAOIILASEAEKAEQINQAGASAVLAKAKAKAEAIRIILAAALTQHNGDAASLTVAEY 282
DB |||||||
223 qaoiilaseAEkAEqINqAGasAVlAkakAKAEaiRIILAAaltQHngdaasltvaeY 282

QY 283 VSAFSLAKDSNTILPSPNGDVTSMVAQAMGYGALTAPVPGRDSSSSRPDOCT 342
DB |||||||
283 vsafSKlAKdsNTilPspngDvtSMvAQaMGyGaltAPvpGRDssSSRPDOct 342

DB 61 vsafsklakdsntilpsnpgdvtsmvagamygaltkapvpgrdpdssssrdovgt 120

QY 343 DAS 345
DB |||
DB 121 das 123

RESULT 10
AAU33158
ID AAU33158 standard; Protein: 222 AA.
XX
XX AAU33158;
AC
DT 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #3649.
DE
XX
XX Human: vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW Immune suppression; immune stimulation; anti-inflammatory; leukaemia.
KW

OY 327 TPDLSGSSRDVQGTDA 345
 ||||||||||||||||
 Db 327 tpdlsissgsrdvqgtas 345

RESULT 7
 AAM78284
 AAM78284 standard; Protein: 306 AA.
 XX
 AC AAM78284;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Fragment of human secreted protein encoded by gene 35.
 XX
 KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 KW
 OS Homo sapiens.
 XX
 PN MO9856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 XX 11-JUN-1998; 98WO-US12125.
 XX
 XX 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058666.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 PR
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DM;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 DR WPI: 1999-080881/07.
 DR N-PSDB: AAX04345.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders
 XX
 XX Disclousre: Page 35; 380pp; English.
 PS
 XX
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a
 CC human immunoglobulin Fc portion (e.g. AAX04302) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAM78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).
 CC
 XX
 XX Sequence 306 AA:
 SQ

Query Match 56.7%; Score 202; DB 20; Length 306;
 Best Local Similarity 100.0%; Pred. No. 2; 6e-189;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 DKVFERESLNASIVDAINQADCGICRLREIKDIHVPPVKESMOMOVEAERKRAT 203
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 94 dkvtrereslnasivdaingdaedcwqjrlrlyeldlnpprvksmqmqaeeatrrat 153

OY 204 VLESEGTRESAINVAEGKKQAOIILASEAKAQINQAGASAVILAKAKAKAATRIILAA 263
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 154 vlesegtresainvaegkkqagqlaseaekeqingagaesaavilakakeatrrilaa 213

OY 264 ALTOHNGDAASLTVAEOYVSFSLKADSNITLIPSPNGVTSNVAQAMGYGALTAP 323
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 214 altqngdaaasltvaeqyvsafsklaxdsntllipnpgdvtsmvadamygaltakp 273

OY 324 VPGTPDLSGSSRDVQGTDA 345
 ||||||||||||||||
 Db 274 vpgtpdlsissgsrdvqgtas 295

RESULT 8
 AAM78566
 AAM78566 standard; Protein: 358 AA.
 XX
 AC AAM78566;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1228.
 XX
 KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 KW
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.

PD 09-AUG-2001.
XX
PF 05-FEB-2001: 2001MO-US03800.
XX
PR 03-FEB-2000: 2000US-0496914.
PR 27-APR-2000: 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT.
XX
XX MPI: 2001-457740/49.
DR N-PSDB: ABA09225.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
XX Claim 20: Page 291: 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC hematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness.
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
XX Sequence 378 AA:
S0

Query Match 61.2%: Score 218: DB 22: Length 378:
Best Local Similarity 99.7%: Pred. No. 6.9e-205:
Matches 318: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Oy 27 RASGCLPNTVAVLPYQOEAHVVERMGRFRLLEGLNLLIPVLDRIKRYVSLKELIVNV 86
|||
Db 49 rassgjlptntvavlfyvgqeaavvermgfhrlllepglllllpvldrikyvsglkelivnv 108
|||
Oy 87 PEQSAVLTLDNNTLOIDGYLYRIMDPKASVGVDEPEYAVNQLAQTTMRSELGKISXKV 146
|||
Db 109 pegsavltldnntloidylyrilmopykasygvdepeyavavtqlatqumse:gklsidkv 168

Oy 147 FRERESLNASTVDAINQADQWGIKRLREIKDIIHVPKPKESNMQOVKERRKRAVLE 206
|||
Db 169 frereslnastvdaingadqwgirclryelkldihvprvkemqmgveeerkratvle 228
|||
Oy 207 SEGTRRESAINVAEKKKQAOILASFAEKAKEQINQAGFSAVLAARAKKAEIRILAAALT 266
|||
Db 229 segtrresainvaeqkkqagilaseakeqingaaesavilakakaeairllaaalt 288
|||
Oy 267 OHNGDAASLTVAEQYVSAFSLKAKDSNTILLPSNPGDVTSMVAQAMGVYALTAKAPVG 326
|||
Db 289 qhngdaaasltvaeqyvsafslakdsntllpsnpgdvtsmvaqangvyaltakapvg 348
|||
Oy 327 TPDSLSSGSSRDVGTDPAS 345
|||
Db 349 tpdslssgsrdvgtldas 367
|||

RESULT 5
AAM79550 standard; Protein: 378 AA.
ID AAM79550:
XX
AC AAM79550:
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3196.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN MO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001: 2001MO-US04098.
XX
PR 03-FEB-2000: 2000US-0496914.
PR 27-APR-2000: 2000US-0560875.
PR 20-JUN-2000: 2000US-0598075.
PR 19-JUL-2000: 2000US-0620325.
PR 01-SEP-2000: 2000US-0654936.
PR 15-SEP-2000: 2000US-0663561.
PR 20-OCT-2000: 2000US-0693325.
PR 30-NOV-2000: 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX MPI: 2001-476283/51.
DR N-PSDB: AAK52683.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 20: Page 281-282: 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

Db 27 rassg1prntvvlfvygqgawwvrmgrfhrllpepnlilpvidrliryvgsiketvlnv 86
 Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
 Db 87 pegsavtlidnvtlqidgylylrmdpykasygvdepeyavtqlagttmrselgklsldkv 146
 Oy 147 FRERESLNASIVDAINQADGCGIRCLRYETKDIHVPRVRESKMOVQEAERKRAATYLE 206
 Db 147 frereslnasivdaingadcgwirlclryetkdlhvprvresmqvqeaerkratyle 206
 Oy 207 SECTRESAINVAEGKKQAOILASEAEKAEQINQAGEASAVLAKAKAEAIRIRIAALT 266
 Db 207 sectrresainvaeqkkqgqilaseekekqinqageasavilakakaeairirilaalt 266
 Oy 267 OHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVQAMGVGALTRAPVPG 326
 Db 267 qhngdaaasltvaeqyvsafslakdsntilpsnpgdvtsmvagmgvgyaltkappg 326
 Oy 327 TPDSLSGSSSRDVGCTDAS 345
 Db 327 tpdsls9gs9srdrvqgldas 345
 RESULT 3
 AAB92659
 ID AAB92659 standard; protein: 356 AA.
 AC AAB92659;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:11010.
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS
 PN EP1074617-A2.
 XX 07-FEB-2001.
 PD
 PF 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ohta T, Sugiyama T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8: SEQ ID 11010; 2537pp + CD-ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 SQ Sequence 356 AA;

Query Match 61.2%; Score 218; DB 22; Length 356;
 Best Local Similarity 99.7%; Pred. No. 6.5e-205;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 RASSGLPRNTVLFVPOGEAWVVERMGRFHRILPEGLNLIPLVDRIRYVQSLKEIVINV 86
 Db 27 rassg1prntvvlfvygqgawwvrmgrfhrllpepnlilpvidrliryvgsiketvlnv 86
 Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
 Db 87 pegsavtlidnvtlqidgylylrmdpykasygvdepeyavtqlagttmrselgklsldkv 146
 Oy 147 FRERESLNASIVDAINQADGCGIRCLRYETKDIHVPRVRESKMOVQEAERKRAATYLE 206
 Db 147 frereslnasivdaingadcgwirlclryetkdlhvprvresmqvqeaerkratyle 206
 Oy 207 SECTRESAINVAEGKKQAOILASEAEKAEQINQAGEASAVLAKAKAEAIRIRIAALT 266
 Db 207 sectrresainvaeqkkqgqilaseekekqinqageasavilakakaeairirilaalt 266
 Oy 267 OHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVQAMGVGALTRAPVPG 326
 Db 267 qhngdaaasltvaeqyvsafslakdsntilpsnpgdvtsmvagmgvgyaltkappg 326
 Oy 327 TPDSLSGSSSRDVGCTDAS 345
 Db 327 tpdsls9gs9srdrvqgldas 345

RESULT 4
 ABB11981
 ID ABB11981 standard; peptide: 378 AA.
 AC ABB11981;
 XX
 DT 11-JAN-2002 (first entry)
 DE Human SLP-2 homologue, SEQ ID NO:2351.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnary; antiulcer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      September 22, 2002, 18:22:47 ; Search time 31.2 Seconds
              (without alignments)
              1267.380 Million cell updates/sec
```

Title:	US-09-898-216-1
Perfect score:	356
Sequence:	1 MLARARGHWGPFAGLGSTG.....RDVQGTDA SXDEELDRVKMS 356

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 segs, 111073796 residues

Word size :

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
```

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Description			
No.	Score	Match	Length	DB	ID
1	354	99.4	356	19	AAW57232
2	218	61.2	356	20	AAW78160
3	218	61.2	356	22	AAW92659
4	218	61.2	378	22	ABB11951
5	218	61.2	378	22	AAW79550
6	218	61.2	3007	22	AAU28053
7	202	56.7	306	20	AAW78284
8	198	55.6	358	22	AAW78566
9	123	34.6	2797	22	AAU33159
10	116	32.6	222	22	AAU33158
11	52	14.6	92	22	AAO09538
					Human integral mem
					Human secreted pro
					Human protein sequ
					Human SLP-2 homolo
					Human protein SEQ
					Novel human secret
					Fragment of human
					Human protein SEQ
					Novel human secret
					Novel human secret
					Human polypeptide

12	50	4.0	463	22	AAU28241	Novel human secret
13	27	7.6	27	20	AAW78286	Fragment of human
14	26	7.3	26	20	AAW78285	Fragment of human
15	26	7.3	26	20	AAW78287	Fragment of human
16	24	6.7	70	20	AAW78288	Fragment of human
17	5.6	6.7	323	22	ABBS59541	Drosophila melanog
18	3.7	3.7	100	21	AAAG50216	Arabidopsis thaliaa
19	3.7	3.7	339	21	AAAG24300	Arabidopsis thaliaa
20	3.7	3.7	358	21	AAAG24299	Arabidopsis thaliaa
21	3.7	3.7	369	21	AAAG24298	Arabidopsis thaliaa
22	13	3.7	382	21	AAAG48817	Arabidopsis thaliaa
23	13	3.7	394	21	AAAY4663	Amino acid sequenc
24	13	3.7	401	21	AAAG48816	Arabidopsis thaliaa
25	13	3.7	411	21	AAAG20661	Arabidopsis thaliaa
26	13	3.7	411	21	AAAG48815	Arabidopsis thaliaa
27	13	3.7	411	21	AAAG50184	Arabidopsis thaliaa
28	13	3.7	515	21	AAAG31395	Arabidopsis thaliaa
29	10	2.8	10	22	AAU25786	Breast cancer-asso
30	8	2.2	15	22	AAU25787	Breast cancer-asso
31	8	2.2	43	22	ABBS7904	Human peptide #355
32	8	2.2	43	22	ABBS3076	Peptide #582 encod
33	8	2.2	43	22	ABBS18544	Protein #543 encod
34	8	2.2	43	22	AAAS3873	Human brain expres
35	8	2.2	43	22	AAAG66261	Human bone marrow
36	8	2.2	43	22	AAAM14131	Peptide #565 encod
37	8	2.2	43	22	AAAM26337	Peptide #574 encod
38	8	2.2	43	22	AAAM01869	Peptide #551 encod
39	8	2.2	504	22	AAAG91264	C glutamicum protee
40	8	2.2	554	22	ABBS71946	Drosophila melanog
41	8	2.2	1015	22	ABGC04137	Novel human diagno
42	8	2.2	1443	22	ABBS68472	Drosophila melanog
43	8	2.2	14	22	AAAB98061	Human peptide #133
44	7	2.0	14	22	AAAB98061	Human peptide #133
45	7	2.0	16	22	AAAB98062	Human peptide #133
46	7	2.0	18	22	AAAB98062	Human peptide #133
47	7	2.0	18	22	AAAB98062	Human peptide #133
48	7	2.0	18	22	AAAB98062	Human peptide #133
49	7	2.0	18	22	AAAB98062	Human peptide #133
50	7	2.0	18	22	AAAB98062	Human peptide #133
51	7	2.0	18	22	AAAB98062	Human peptide #133
52	7	2.0	18	22	AAAB98062	Human peptide #133
53	7	2.0	18	22	AAAB98062	Human peptide #133
54	7	2.0	18	22	AAAB98062	Human peptide #133
55	7	2.0	18	22	AAAB98062	Human peptide #133
56	7	2.0	18	22	AAAB98062	Human peptide #133
57	7	2.0	18	22	AAAB98062	Human peptide #133
58	7	2.0	18	22	AAAB98062	Human peptide #133
59	7	2.0	18	22	AAAB98062	Human peptide #133
60	7	2.0	18	22	AAAB98062	Human peptide #133
61	7	2.0	18	22	AAAB98062	Human peptide #133
62	7	2.0	18	22	AAAB98062	Human peptide #133
63	7	2.0	18	22	AAAB98062	Human peptide #133
64	7	2.0	18	22	AAAB98062	Human peptide #133
65	7	2.0	18	22	AAAB98062	Human peptide #133
66	7	2.0	18	22	AAAB98062	Human peptide #133
67	7	2.0	18	22	AAAB98062	Human peptide #133
68	7	2.0	18	22	AAAB98062	Human peptide #133
69	7	2.0	18	22	AAAB98062	Human peptide #133
70	7	2.0	18	22	AAAB98062	Human peptide #133
71	7	2.0	18	22	AAAB98062	Human peptide #133
72	7	2.0	18	22	AAAB98062	Human peptide #133
73	7	2.0	18	22	AAAB98062	Human peptide #133
74	7	2.0	18	22	AAAB98062	Human peptide #133
75	7	2.0	18	22	AAAB98062	Human peptide #133
76	7	2.0	18	22	AAAB98062	Human peptide #133
77	7	2.0	18	22	AAAB98062	Human peptide #133
78	7	2.0	18	22	AAAB98062	Human peptide #133
79	7	2.0	18	22	AAAB98062	Human peptide #133
80	7	2.0	18	22	AAAB98062	Human peptide #133
81	7	2.0	18	22	AAAB98062	Human peptide #133
82	7	2.0	18	22	AAAB98062	Human peptide #133
83	7	2.0	18	22	AAAB98062	Human peptide #133
84	7	2.0	18	22	AAAB98062	Human peptide #133
85	7	2.0	18	22	AAAB98062	Human peptide #133
86	7	2.0	18	22	AAAB98062	Human peptide #133
87	7	2.0	18	22	AAAB98062	Human peptide #133
88	7	2.0	18	22	AAAB98062	Human peptide #133
89	7	2.0	18	22	AAAB98062	Human peptide #133
90	7	2.0	18	22	AAAB98062	Human peptide #133
91	7	2.0	18	22	AAAB98062	Human peptide #133
92	7	2.0	18	22	AAAB98062	Human peptide #133
93	7	2.0	18	22	AAAB98062	Human peptide #133
94	7	2.0	18	22	AAAB98062	Human peptide #133
95	7	2.0	18	22	AAAB98062	Human peptide #133
96	7	2.0	18	22	AAAB98062	Human peptide #133
97	7	2.0	18	22	AAAB98062	Human peptide #133
98	7	2.0	18	22	AAAB98062	Human peptide #133
99	7	2.0	18	22	AAAB98062	Human peptide #133
100	7	2.0	18	22	AAAB98062	Human peptide #133

ALIGNMENTS

RESULT	1
AAW57232	
ID	AAW57232 standard; Protein; 356 AA.
XX	
AC	AAW57232;
XX	
DT	03-AUG-1998 (first entry)
XX	
DE	Human integral membrane protein.
XX	
KW	Human; integral membrane protein; IMP; cancer; anaemia; prostate;
KW	breast; pancreatic; tumour; ion transport.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Misc-difference 143
FT	/label= unknown
FT	/note= "encoded by NTG"
FT	Misc-difference 173
FT	/note= "encoded by CTN"
FT	Misc-difference 346
FT	/label= unknown
FT	/note= "encoded by NTT"
XX	
UN	US5763589-A.
XX	
PD	09-JUN-1998.
XX	
PE	09-JAN-1997; 97US-0781562.
XX	
PR	09-JAN-1997; 97US-0781562.
XX	
PA	(INCY-) INCYTE PHARM INC.

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035602; CAB8270.1; -
DR EMBL; AL161571; CAB81408.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART; SM00244; PHB; 1.
KW Hypothetical protein.
SQ SEQUENCE 515 AA; 55923 MW; 420FE3DB5B7AFD8A CRC64;

Query Match 3.7%; Score 13; DB 10; Length 515;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 TOLAOTMRSELC 140
DB 153 TOLAOTMRSELC 165

RESULT 13
O9ZDKO PRELIMINARY: PRT: 311 AA.

AC O9ZDKO
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 34.0 KDA PROTEIN.
GN RP328.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_Taxid=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAURID E;
RA MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of mitochondria."
RL Nature 396:133-140(1998).
DR EMBL; AJ235271; CAI4788.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART; SM00244; PHB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 311 AA; 34040 MW; F74F625EC567E5CA CRC64;

Query Match 3.1%; Score 11; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAOTMRSE 137
DB 109 TOLAOTMRSE 119

RESULT 14
O92IG8 PRELIMINARY: PRT: 312 AA.

AC O92IG8
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN RC0452.
GN RC0452.
OS Rickettsia conorii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_Taxid=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
RL Science 293:2093-2098(2001).
DR EMBL; AE008609; AL102990.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 312 AA; 34025 MW; A41878CANAABC49F5 CRC64;

Query Match 3.1%; Score 11; DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAOTMRSE 137
DB 109 TOLAOTMRSE 119

RESULT 15
O60121 PRELIMINARY: PRT: 354 AA.

AC O60121
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 39.3 KDA PROTEIN C1665.07C IN CHROMOSOME II.
GN SPIC1665.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Beck A.,
RA Reinhardt R.;
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
DR EMBL; AL023554; CAI19027.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR SMART; SM00244; PHB; 1.
DR PROSITE; PS01270; BAND_7; FALSE_NEG.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39274 MW; 66B09E6A12BDC030 CRC64;

Query Match 3.1%; Score 11; DB 3; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 DPKASYGVED 121
DB 126 DPKASYGVED 136

Search completed: September 22, 2002, 18:36:37
Job time: 294 sec

RT "prohibitins, Stomatins, and plant Disease Response Genes Comprise a
RT Protein Superfamily That Controls Cell Proliferation, Ion Channel
RT Regulation, and Death."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF236372; AAF68388.1; -
DR InterPro: IPR001107; Band_7
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 394 AA: 43277 MW: 52A409DCFI5C5A45 CRC64:

Query Match 3.7%; Score 13; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 148 QLAQTTMRSELGK 160

RESULT 10
O9LW0 PRELIMINARY: PRT; 401 AA.
AC O9LW0:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE SIMILARITY TO STOMATIN LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and YAC
RT clones."
RL DNA Res. 7:31-63(2000).
DR EMBL: AB018115; BAA97132.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 401 AA: 43607 MW: 28E8BD732A370AB6 CRC64:

Query Match 3.7%; Score 13; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 196 QLAQTTMRSELGK 208

RESULT 11
O93VP9 PRELIMINARY: PRT; 411 AA.
AC O93VP9:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE HYPOTHETICAL 45.0 KDA PROTEIN.
GN AT4G27580.
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G27580 (GI:7269612).";
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G27580 (GI:7269612).";
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY059109; AAL15215.1; -
DR EMBL: AY034924; AAK59431.1; -
KW Hypothetical protein.
SQ SEQUENCE 411 AA: 45020 MW: 94E24ED4C9B89633 CRC64:

Query Match 3.7%; Score 13; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 153 QLAQTTMRSELGK 165

RESULT 12
O9T082 PRELIMINARY: PRT; 515 AA.
AC O9T082:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE HYPOTHETICAL 55.9 KDA PROTEIN.
GN T29A15.70 OR AT4G27580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
RA Volckaert G., Hohnel J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robben J., Grymoprez B., Volckaert G., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

Qy	157	IVDAINQADDCGIRCLREIKDHPVPRKESQMOYEAERKRATVLESCTRESAIN	216
Db	157	IVDAINQADDCGIRCLREIKDHPVPRKESQMOYEAERKRATVLESCTRESAIN	216
Qy	217	VAEGKQAOIILASEAKAEQIINQAGSAVLAKKAKAEAIRILIA	262
Db	217	VAEGKQAOIILASEAKAEQIINQAGSAVLAKKAKAEAIRILIA	262
RESULT	7		
ID	09W1E7	PRELIMINARY;	PRT: 323 AA.
AC	09W1E7		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	CG2970	PROTEIN.	
GN	CG2970		
OS	Drosophila melanogaster (fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_taxid=7227;		
NP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Hote R.A., Evans C.A., Gocayne J.D.,		
RA	Amantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Burton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.H.C., Blazel R.G., Champe M., Pfaffner B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abdill J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Bessley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Bortova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov J.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Garfield A.E., Gay N.S., Gelbart W.M., Glasser K.,		
RA	Giodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ibeguam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J.Z., Li Z., Liang Y., Lin X.,		
RA	Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry G., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy N., Murphy K., Nusskern D.R., Pacleb J.M.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spitskas R., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Wang Z.-Y., Massaman D., Weller L., Venter E., Wang A.H., Wang X.,		
RA	Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster."		
RL	Science 287:2185-2195(2000).		
DR	EMBL: AEO03462; AAF47110.1; -		
DR	FlyBase: FBgn0034936; CG2970.		
DR	InterPro: IPR001107; Band 7.		
DR	InterPro: IPR001972; Stomatin.		
DR	Pfam: PF01145; Band 7; 1.		

```

DR PRINTS: PR00721; STOMATIN.
DR SMART: SMO0244; PHB; 1.
SQ SEQUENCE 323 AA; 35668 MW; D5C6241445FEF4DE CRC64;

Query Match          5.6%; Score 20; DB 5; Length 323;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      40 FVPOGEAMVVERMGFRHRL 59
        |||||||
DB       2 FVPOGEAMVVERMGFRHRL 21

RESULT      8
O9XVP9
ID O9XVP9 PRELIMINARY; PRT; 334 AA.
AC O9XVP9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F30A10.5 PROTEIN.
GN F30A10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (CCr-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81072; CAB03018.1; .
DR InterPro; IPR001107; Band_7.
DR InterPro; IPR001972; StomatIn.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SMO0244; PHB; 1.
SQ SEQUENCE 334 AA; 36704 MW; 10F98BB9AB44E5ED CRC64;

Query Match          3.9%; Score 14; DB 5; Length 334;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY      40 FVPOGEAMVVERMG 53
        |||||||
DB       47 FVPOGEAMVVERMG 60

RESULT      9
O9M585
ID O9M585 PRELIMINARY; PRT; 394 AA.
AC O9M585;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE STOMATIN-LIKE PROTEIN.
GN STM1.
OS zea mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Nadimpalli R., Yalpani N., Johal G.S., Simmons C.R.;
```

```
RESULT 4
ID 09P042 PRELIMINARY: PRT: 342 AA.
AC 09P042:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HSP108.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxId=9606;
RN 11)
RP SEQUENCE FROM N.A.
RC TISSUE=BLD;
RX MEDLINE-20499367; PubMed-11042152;
RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,
RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;
RT Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.
RL Genome Res. 10:1546-1560(2000).
DR EMBL: AF161458; AAF29073.1; -.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 342 AA; 37145 MW; 79580E3BDE53E5 CRC64;

Query Match 40.2%; Score 143; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e-133;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 TVLESGTRESAINVAEGKKQAQIILASEAEKAEQINQAGEASAVLAKAKAEAIRILA 262
DB 189 TVLESGTRESAINVAEGKKQAQIILASEAEKAEQINQAGEASAVLAKAKAEAIRILA 248
QY 263 AALTOHNGAASLTYAEQVSAFSLKADSNITLLPSNPGDVTSMVAQAMGYGALTGA 322
DB 249 AALTOHNGAASLTYAEQVSAFSLKADSNITLLPSNPGDVTSMVAQAMGYGALTGA 308
QY 323 PVPGTPDSLSSGSSRDVGCTDAS 345
DB 309 PVPGTPDSLSSGSSRDVGCTDAS 331

RESULT 5
ID 09DCG8 PRELIMINARY: PRT: 353 AA.
AC 09DCG8:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 0610038F01RIK PROTEIN.
GN 0610038F01RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE-21083660; PubMed-11217851;
RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arawata T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nixaldo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib L.F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bash G.,
RA Lake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK002793; BAB22363.1; -.
DR MGD: MGI:1913842; 0610038F01RIK.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 353 AA; 38530 MW; D317CBB7E32F8863 CRC64;
```

```
Query Match 29.8%; Score 106; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.6e-96;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 157 IYDAINQADQNCIRCLRREIKIHPPRYKESMOMQVVEARRKRTVLESEGTRESAIN 216
DB 157 IYDAINQADQNCIRCLRREIKIHPPRYKESMOMQVVEARRKRTVLESEGTRESAIN 216
QY 217 VAEGRKQAQIILASEAEKAEQINQAGEASAVLAKAKAEAIRILA 262
DB 217 VAEGRKQAQIILASEAEKAEQINQAGEASAVLAKAKAEAIRILA 262

RESULT 6
ID 099JB2 PRELIMINARY: PRT: 353 AA.
AC 099JB2:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RIKEN CDM 0610038F01 GENE (STOMATIN-LIKE PROTEIN 2).
GN 0610038F01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RN 11)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RC STRAIN=FVB; TISSUE=KIDNEY;
RP SEQUENCE FROM N.A.
RA Chang J.G., Chan W.L.;
RT "Mouse stomatin-like protein 2 (MSLP2) mRNA.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC003425; AAH03425.1; -.
DR EMBL: AF323178; AAG53404.1; -.
DR MGD: MGI:1913842; 0610038F01RIK.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 353 AA; 38385 MW; 391D269576F566BB CRC64;
```

```
Query Match 29.8%; Score 106; DB 11; Length 353;
```

RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Nakamura Y., Nagahara K., Masuho Y.,
 RA Niimura K., Iwayanagi T.,
 RT "NDO human cDNA sequencing project."
 RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS, EPITHELIOID CARCINOMA;
 RA Strausberg R.;
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF10167; AAF09142.1; -
 DR EMBL: AF282596; AAF91466.1; -
 DR EMBL: BC003025; AAH03025.1; -
 DR EMBL: BC002442; AAH02442.1; -
 DR EMBL: AK027405; BAB5091.1; -
 DR EMBL: BC014990; AAH14990.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 356 AA; 38534 MW; 672331B57C82654E CRC64;

Query Match 61.2%; Score 218; DB 4; Length 356;
 Best Local Similarity 99.7%; Pred. No. 1.7e-206;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 RASSGLPRTVVLFVPOEAMVVERKGRFRILEPGLNILEFVLRIRYQSLKEIVNV 86
 DB 27 RASSGLPRTVVLFVPOEAMVVERKGRFRILEPGLNILEFVLRIRYQSLKEIVNV 86
 QY 87 PQSAVTLNDVTLQIDGVLYLRIMPYKASVGEDEPEYAVTQLAOTTMSELSGXDKV 146
 DB 87 PQSAVTLNDVTLQIDGVLYLRIMPYKASVGEDEPEYAVTQLAOTTMSELSGXDKV 146
 QY 147 FFRERSLNASYDAINQAADCWGIRCLREIKDIHYPPRVKESMOMQVAEERKRATVLE 206
 DB 147 FFRERSLNASYDAINQAADCWGIRCLREIKDIHYPPRVKESMOMQVAEERKRATVLE 206
 QY 207 SEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKAKAERIRILAAALT 266
 DB 207 SEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKAKAERIRILAAALT 266
 QY 267 QHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTKAPVPG 326
 DB 267 QHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTKAPVPG 326
 QY 327 TPDLSLSSGSSRDVQGTAS 345
 DB 327 TPDLSLSSGSSRDVQGTAS 345

RESULT 2
 Q96FY2 PRELIMINARY; PRT; 356 AA.
 AC Q96FY2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE STOMATIN-LIKE 2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: BC010152; AAH10152.1; -
 SQ SEQUENCE 356 AA; 38518 MW; 6FC0BF899E1919A7 CRC64;
 Query Match 56.7%; Score 202; DB 4; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1e-190;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 DKVFRERSLNASYDAINQAADCWGIRCLREIKDIHYPPRVKESMOMQVAEERKRAT 203
 DB 144 DKVFRERSLNASYDAINQAADCWGIRCLREIKDIHYPPRVKESMOMQVAEERKRAT 203
 QY 204 VLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKAKAERIRILAA 263
 DB 204 VLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKAKAERIRILAA 263
 QY 264 ALTOHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTKAP 323
 DB 264 ALTOHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTKAP 323
 QY 324 VPGTPDLSLSSGSSRDVQGTAS 345
 DB 324 VPGTPDLSLSSGSSRDVQGTAS 345

RESULT 3
 Q60376 PRELIMINARY; PRT; 357 AA.
 AC Q60376;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PL11659.4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FIBROBLAST OR FORESKIN;
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
 RA Burthart-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Ganes J., Dangnan L., Poundstone P.,
 RA Christensen M., Georges A., Avila J., Liu S., Attix C., Andreise T.,
 RA Trankheim M., Amico-Keller G., Coefficient J., Duarte S., Lucas S.,
 RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
 RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
 RA Carrano A.V.;
 RT "Sequence analysis of a human P1 clone containing the XRC9 DNA repair
 RT gene."
 RL Submitted (Mar-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC004472; AAC07983.1; -
 DR InterPro: IPR001107; Band_7.
 DR Pfam: PF01145; Band_7; 1.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 357 AA; 38749 MW; 512632B83028135A CRC64;

Query Match 42.7%; Score 152; DB 4; Length 357;
 Best Local Similarity 100.0%; Pred. No. 2.3e-141;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 VEAERKRATVLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKA 253
 DB 194 VEAERKRATVLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAAGEASAVLAKAKA 254
 QY 254 KAEAIRILAAALTOHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAM 313
 DB 254 KAEAIRILAAALTOHNGDAASLTVAEOYVSFAFKLAKDSNTILLPSNPGDVTSMVAQAM 314
 QY 314 GYGALTKAPVPGTPDLSLSSGSSRDVQGTAS 345
 DB 315 GYGALTKAPVPGTPDLSLSSGSSRDVQGTAS 346

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:31:43 : Search time 25.31 Seconds
(without alignments)
2433.275 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLARARCHMGCPFAEGLSTG.....RDVGTDPASXDELDPRVKMS 356

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
562222 seqs, 172994929 residues

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218	61.2	356	4	Q9UJZ1
2	202	56.7	356	4	Q96FY2
3	152	42.7	357	4	O60376
4	143	40.2	342	4	O9P042
5	106	29.8	353	11	Q9DC68
6	106	29.8	353	11	Q99J82
7	20	5.6	323	5	O9W1F7
8	14	3.9	334	5	O9XVP9
9	13	3.7	394	10	Q9M585
10	13	3.7	401	10	Q9LWV0
11	13	3.7	411	10	Q93VP9
12	13	3.7	515	10	Q9T082
13	11	3.1	311	16	Q9ZDK0
14	11	3.1	312	16	O92IC8
15	11	3.1	354	3	O60121
16	9	2.5	357	5	O9U4M5

17	9	2.5	374	5	077376	077376 plasmodium
18	8	2.2	235	9	Q94M69	Q94M69 streptococc
19	8	2.2	248	16	Q9ABD4	Q9ABD4 caulobacter
20	8	2.2	263	16	Q9HXE1	Q9HXE1 pseudomonas
21	8	2.2	398	10	O04471	O04471 arabisopsis
22	8	2.2	477	5	O62022	O62022 caenorhabdi
23	8	2.2	523	5	O77238	O77238 drosophila
24	8	2.2	554	5	O9VWX7	O9VWX7 drosophila
25	8	2.2	658	5	O18703	O18703 caenorhabdi
26	8	2.2	716	5	O80256	O80256 vibrio chol
27	8	2.2	754	5	Q9NED3	Q9NED3 leishmania
28	8	2.2	1443	5	Q9VPR0	Q9VPR0 drosophila
29	8	2.2	8817	2	O53840	O53840 polyangium
30	7	2.0	47	6	Q28882	Q28882 ovis aries
31	7	2.0	95	5	Q26830	Q26830 trypanosoma
32	7	2.0	95	16	Q91054	Q91054 pseudomonas
33	7	2.0	110	2	Q9APJ7	Q9APJ7 hyphomicrob
34	7	2.0	112	16	Q26072	Q26072 helicobacte
35	7	2.0	112	16	Q92J64	Q92J64 helicobacte
36	7	2.0	120	5	Q26834	Q26834 trypanosoma
37	7	2.0	121	17	Q9YC53	Q9YC53 aeropyrum p
38	7	2.0	129	16	Q98AP9	Q98AP9 rhizobium l
39	7	2.0	133	2	Q939R0	Q939R0 fibrobacter
40	7	2.0	134	11	O88728	O88728 mus musculu
41	7	2.0	135	5	Q9TXE2	Q9TXE2 chlamys nlp
42	7	2.0	135	17	Q9YB38	Q9YB38 aeropyrum p
43	7	2.0	136	16	Q92MW3	Q92MW3 rhizobium m
44	7	2.0	137	15	Q91HW7	Q91HW7 human immun
45	7	2.0	142	11	Q9CS76	Q9CS76 mus musculu

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	356 AA.
1	Q9UJZ1			
Q9UJZ1	Q9UJZ1	PRELIMINARY:	PRT:	356 AA.
AC	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	MEMBRANE ASSOCIATED PROTEIN SLP-2 (STOMATIN-LIKE PROTEIN 2) (STOMATIN-LIKE 2) (CDNA FLJ14499 FIS, CLONE NTZRM100080, WEAKLY SIMILAR TO UNC-1 PROTEIN) (STOMATIN-LIKE 2).			
DE	UNC-1 PROTEIN) (STOMATIN-LIKE 2).			
GN	HUSLP2 OR SLP2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=HEART MUSCLE;			
RA	Wang Y., Morrow J.S.;			
RT	"Identification and Characterization of Human SLP-2, a Novel Homologue of Stomatlin (band 7.2b) Present in Erythrocytes and Other Tissues.";			
RL	Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=BRIN;			
RA	Owczarek C.M., Treutlein H.R., Portbury K.J., Gulluyan L.M., Kola I., Hertzog P.J.;			
RT	"A novel member of the stomatin/EPB72/MC-2 family, stomatin-like protein 2 (SLP2), is ubiquitously expressed and localizes to HSA chromosome 9p13.1.";			
RT	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=JUNG CARCINOMA, SKIN, AND MELANOTIC MELANOMA.;			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Isoagai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			

QY 182 -----VPPRVKESM----- 190
Db 132 TGVAAERDAGIRAEACCKEMLDVKMDPTKIADSKRAPELOKSAFSEVINIKATAQOLA 191
QY 191 -----OMOVAERRRKRAVLE-----SEGTRESAINV 217
Db 192 YELGAREEOQIKROEIEIEVYQKKOIAVEAOEILRTDKELIATVRRPAFAEAHRIQOI 251
QY 218 AEGKKOAOILASAEKAEQINOAGASAVLAKAKAKAEAIRI-LAALTCCHNGDAASL 276
Db 252 AEGEKYKQVLAQAE-AEKIRK-IGEAFAVTEAMGKAERKKLAEAYQKYGDAAKMA 309
QY 277 TVAEQYVSASFSLA-----KDSNTILLPSNPGDYSMAVQAAMGVYGALTKAPVPGTPDLS 332
Db 310 LVLEALPQIAAKIAAPLTVDELIVLSGDSKVTSEVNRLL-----AELPASVHALT 361
QY 333 SGSSRDVGCTDAS 345
Db 362 -----GVDLS 366

RESULT 8
US-08-762-106-8
; Sequence 8, Application US/08762106
; Patent No. 5948677
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: READING FRAME INDEPENDENT EPTIOTE
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harris Brotman
; STREET: 202 Coast Blvd., Suite 111
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,106
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brotman, Harris F.
; REGISTRATION NUMBER: 35,461
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 654-2428
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-762-106-8

Query Match 5.9%; Score 104.5; DB 2; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

QY 55 FHRILEPGLNIIPLVDRII--YVOSLKEIVINVPQSAVTLQIDGVLRLIMDP 112
Db 257 FDLTVSSGY-IYDPVMEVETAFMPMLKEQAIGYLAR-----GVVARRVVDK 302
QY 113 YKASYGVEDEPVAVTQLAQTMTSELGKLSXKDVFRERESLNSIYDAINQAADCWIRC 172
Db 303 L-----VED-----AAALAAAN--RSTLA-----DKAASTAAATYDA-----WAERQ 336

QY 173 LRYEIKDIHVPVRKESMOMQVEERRKRAVLE-----SEGTRESAINVAEGKKOQ 225
Db 337 AKME-----AELQGELEFAVRRPTFVLRBLKPAVASADAVPA--AAELTQAE 364
QY 226 ILAS---EAEKAEQINOAGASAVLAKAKAKAEAIRIILAAALTOHNGDAASLTVAEOY 282
Db 385 EAANNAKWEADKAFAEKAFAEAEEQKALLEELATATAAEEERGEPPAEPSPS--- 441
QY 283 VSAFSKLAKNSNTILPS--NPGDYISMVAQAAMGVYGALTKAPVPGTPD 329
Db 442 -----LPDGVPEVDVEAEVAKAV--EAVKPPVKEVTD 472

RESULT 9
US-08-745-404-2
; Sequence 2, Application US/08745404B
; Patent No. 6096717
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: Method for Producing Tagged Genes,
; FILE REFERENCE: 2087-961422
; CURRENT APPLICATION NUMBER: US/08/745,404B
; EARLIER FILING DATE: 1996-11-08
; EARLIER APPLICATION NUMBER: 08/000,619
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 516
; TYPE: PRP
; ORGANISM: Chlamydomonas
; US-08-745-404-2

Query Match 5.9%; Score 104.5; DB 3; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

QY 55 FHRILEPGLNIIPLVDRII--YVOSLKEIVINVPQSAVTLQIDGVLRLIMDP 112
Db 257 FDLTVSSGY-IYDPVMEVETAFMPMLKEQAIGYLAR-----GVVARRVVDK 302
QY 113 YKASYGVEDEPVAVTQLAQTMTSELGKLSXKDVFRERESLNSIYDAINQAADCWIRC 172
Db 303 L-----VED-----AAALAAAN--RSTLA-----DKAASTAAATYDA-----WAERQ 336
QY 173 LRYEIKDIHVPVRKESMOMQVEERRKRAVLE-----SEGTRESAINVAEGKKOQ 225
Db 337 AKME-----AELQGELEFAVRRPTFVLRBLKPAVASADAVPA--AAELTQAE 364
QY 226 ILAS---EAEKAEQINOAGASAVLAKAKAKAEAIRIILAAALTOHNGDAASLTVAEOY 282
Db 385 EAANNAKWEADKAFAEKAFAEAEEQKALLEELATATAAEEERGEPPAEPSPS--- 441
QY 283 VSAFSKLAKNSNTILPS--NPGDYISMVAQAAMGVYGALTKAPVPGTPD 329
Db 442 -----LPDGVPEVDVEAEVAKAV--EAVKPPVKEVTD 472

RESULT 10
US-09-320-774-8
; Sequence 8, Application US/09320774
; Patent No. 6265545
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: READING FRAME INDEPENDENT EPTIOTE
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harris Brotman
; STREET: 202 Coast Blvd., Suite 111

CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/320,774
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/762,106
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brotman, Harris F.
REGISTRATION NUMBER: 35,461
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 654-2428
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-320-774-8

Query Match 5.9%; Score 104.5; DB 4; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

OY 55 FHRIIEPGNIIIPVDRIR--YVQSLKEIYINVPQSAVTLDNVTLOIDGVLVLRIMP 112
DB 255 FRLVSSG-IYDPVAREVETAFMPWLKEQALGYLAR-----GVYARRVVK 302
OY 113 YKASYGVDEPEYAVTQLOATMRSELGKLSXDKVFERESLNASIVDAINQADCGIRC 172
DB 303 L-----VED-----AAALAN--RSTLA-----DKAASATAVDA-----WAERQ 336
OY 173 LYEIKDHIHVPKVESMOMQYEAERKRATYLE-----SEGTRFSINVAEGKKQO 225
DB 337 AKME-----AELOGKELEAVARRPTFVLRLEKPAVASADAVEA--AAELTQAE 384
OY 226 ILAS---EAEKAEQIINQAGEASAVLAKAKAEAIRIILAAALTOHNGDAASLTVAEQY 282
DB 385 EAAAKWEADKAEAEKAEAEAAAEOKALLEELAAATAAEAEERGEPPAEPPS--- 441
OY 283 VSAFSKLAKDSNTILPS--NPGDYTSMAVAQMGVYGALTAKAPVPGTPD 329
DB 442 -----LPGDVEPVDAEAVAKAV---EAVPKPPVKVTD 472

RESULT 11
US-08-762-106-9
Sequence 9, Application US/08762106
Patent No. 5948677
GENERAL INFORMATION:
APPLICANT: Jarvik, Jonathan W.
TITLE OF INVENTION: READING FRAME INDEPENDENT EPITOPE
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harris Brotman
STREET: 202 Coast Blvd., Suite 111
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,106
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brotman, Harris F.
REGISTRATION NUMBER: 35,461
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 654-2428
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-762-106-9

Query Match 5.9%; Score 104.5; DB 2; Length 527;
Best Local Similarity 25.6%; Pred. No. 0.069;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

OY 55 FHRIIEPGNIIIPVDRIR--YVQSLKEIYINVPQSAVTLDNVTLOIDGVLVLRIMP 112
DB 268 FRLVSSG-IYDPVAREVETAFMPWLKEQALGYLAR-----GVYARRVVK 313
OY 113 YKASYGVDEPEYAVTQLOATMRSELGKLSXDKVFERESLNASIVDAINQADCGIRC 172
DB 314 L-----VED-----AAALAN--RSTLA-----DKAASATAVDA-----WAERQ 347
OY 173 LYEIKDHIHVPKVESMOMQYEAERKRATYLE-----SEGTRFSINVAEGKKQO 225
DB 348 AKME-----AELOGKELEAVARRPTFVLRLEKPAVASADAVEA--AAELTQAE 395
OY 226 ILAS---EAEKAEQIINQAGEASAVLAKAKAEAIRIILAAALTOHNGDAASLTVAEQY 282
DB 396 EAAAKWEADKAEAEKAEAEAAAEOKALLEELAAATAAEAEERGEPPAEPPS--- 452
OY 283 VSAFSKLAKDSNTILPS--NPGDYTSMAVAQMGVYGALTAKAPVPGTPD 329
DB 453 -----LPGDVEPVDAEAVAKAV---EAVPKPPVKVTD 483

RESULT 12
US-09-320-774-9
Sequence 9, Application US/09320774
Patent No. 6265545
GENERAL INFORMATION:
APPLICANT: Jarvik, Jonathan W.
TITLE OF INVENTION: READING FRAME INDEPENDENT EPITOPE
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harris Brotman
STREET: 202 Coast Blvd., Suite 111
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/320,774
FILING DATE:
CLASSIFICATION:

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington, D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/216,894
;; FILING DATE: 24-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 564 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-216-894-2

Query Match 5.6%; Score 99.5; DB 2; Length 564;
Best Local Similarity 23.5%; Pred. No. 0.24;
Matches 50; Conservative 42; Mismatches 80; Indels 41; Gaps 7;
QY 113 YKASYGVED-----PEYATQLAQ-----TTMRSELGLSXDKVFRERESLNASIYD 159
DB 206 WQATFEGGHPKPSDLVPRGSPSOLQQAENNITNSKREMTKLEKYKKAKEKEL---D 260
QY 160 AINQAADC-----WGIRCLRYEIKDIIHVPVKESMOMQVEAERR 199
DB 261 AINRAATKLEENNOAVKAAHKAEEKAKTFFORLITFESENINLKKRPNDVAVSNRDKKNS 320
QY 200 KRATVLESEGTRES-AINVAEGKKQAQILLASEEKAQINOAGASAVLAKAKAKAEAI 258
DB 321 ETAKTDEVKORAAEAKAVETEKORAAETKYVAEAK-RKAAEAAKAVETEKORAAEAT 379
QY 259 RILAAALTOHNGDAASLTVAEQVSAFSKLA 291
DB 380 KY-AEAEKOKAAEAAKAVETEKORAAEATKVAE 411

Search completed: September 22, 2002, 18:21:40
Job time: 3406 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 22, 2002, 18:14:17 ; Search time 196.92 Seconds
(without alignments)
636.325 Million cell updates/sec

Title: US-09-898-216-1
Perfect score: 1767
Sequence: 1 MLARARGHWGPFAGLSTG.....RDVQGTDA SXDELDREVKMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1:	/cgn2_6/ptodata/1/paa/PTUS.COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US082.COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US083.COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US084.COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US085.COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US086.COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US087.COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US088.COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US089.COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US090.COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US091.COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US092.COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US093.COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US094.COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US095.COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US096.COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US097.COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US098.COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US099.COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US100.COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US101.COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US102.COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US103.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1646	93.2	356	1 PCT-US98-12125-145	Sequence 145, App
2	1646	93.2	356	16 US-09-209-462B-150	Sequence 150, App
3	1646	93.2	378	1 PCT-US01-03800A-2351	Sequence 2351, App
4	1646	93.2	378	1 PCT-US01-04098A-3196	Sequence 3196, App
5	1643	93.0	3007	1 PCT-US01-04942A-222	Sequence 222, App
6	1643	93.0	3019	1 PCT-US01-0856-10594	Sequence 10594, App
7	1635	92.5	358	1 PCT-US01-04098A-1228	Sequence 1228, App

8	1493	84.5	306	1 PCT-US98-12125-272	Sequence 272, App
9	1493	84.5	306	16 US-09-209-462B-411	Sequence 411, App
10	1488	84.2	305	16 US-09-298-733-36	Sequence 36, App
11	1488	84.2	305	16 US-09-298-733A-36	Sequence 36, App
12	1488	84.2	305	21 US-09-723-594-36	Sequence 36, App
13	1488	84.2	305	21 US-09-724-497-36	Sequence 36, App
14	1064.5	60.2	323	20 US-09-614-150-5415	Sequence 5415, App
15	1064.5	60.2	323	26 US-60-191-637-5433	Sequence 5433, App
16	946	53.5	463	1 PCT-US01-04942A-598	Sequence 598, App
17	894.5	50.6	753	26 US-60-167-217-5529	Sequence 5529, App
18	826	46.7	411	20 US-09-620-394B-4030	Sequence 4030, App
19	826	46.7	411	21 US-09-708-427-12542	Sequence 12542, App
20	803	45.4	503	21 US-09-708-427-56695	Sequence 56695, App
21	799	45.2	356	21 US-09-708-427-56696	Sequence 56696, App
22	796	45.0	344	21 US-09-708-427-56697	Sequence 56697, App
23	794.5	45.0	394	17 US-09-395-197-2	Sequence 2, App
24	794.5	45.0	394	21 US-09-767-129-2	Sequence 2, App
25	794.5	45.0	436	21 US-09-708-427-57092	Sequence 57092, App
26	791.5	44.8	349	21 US-09-708-427-57093	Sequence 57093, App
27	716	40.5	315	22 US-09-806-866-13	Sequence 13, App
28	716	40.5	315	22 US-09-806-866-15	Sequence 15, App
29	712.5	40.3	315	22 US-09-806-866-17	Sequence 17, App
30	711.5	40.3	316	22 US-09-806-866-11	Sequence 11, App
31	661.5	37.4	275	26 US-60-324-109-32115	Sequence 32115, App
32	619.5	35.1	263	18 US-09-417-507-32569	Sequence 32569, App
33	588	33.3	283	18 US-09-708-427-57094	Sequence 57094, App
34	584.5	33.1	265	21 US-09-620-394B-4031	Sequence 4031, App
35	534.5	30.2	253	20 US-09-708-427-12543	Sequence 12543, App
36	534.5	30.2	253	21 US-09-248-796-16859	Sequence 16859, App
37	480.5	27.2	279	16 US-09-543-681A-6978	Sequence 6978, App
38	480.5	27.2	320	19 US-09-489-309A-13396	Sequence 13396, App
39	477.5	27.0	320	19 US-09-489-309A-13396	Sequence 13396, App
40	474	26.8	18	18 US-09-897-516-7262	Sequence 7262, App
41	471.5	26.7	286	26 US-60-215-161-7262	Sequence 7262, App
42	471.5	26.7	286	26 US-09-739-449-11515	Sequence 11515, App
43	465	26.3	349	22 US-09-803-110-11515	Sequence 11515, App
44	465	26.3	349	22 US-09-902-540-11334	Sequence 11334, App
45	460	26.0	368	23	

ALIGNMENTS

RESULT: 1
PCT-US98-12125-145
Sequence 145, Application PC/TUS9812125
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
NUMBER OF SEQUENCES: 318
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12125
FILING DATE: Jan 01, 1990
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P2008complete

Best Local Similarity 95.0%; Pred. No. 2e-141;
Matches 340; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

```
OY 1 MLARAARGHMGFPABG--LSTGFWPRSGRASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
OY 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178
DB 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178
OY 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238
DB 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238
OY 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEOYVSAFSKLAKDSNTILL 298
DB 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEOYVSAFSKLAKDSNTILL 298
OY 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOGTDASXDEIDRYKMT 356
DB 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOGTDASXDEIDRYKMT 356
```

RESULT 6

```
PCT-US01-08656-10594
; Sequence 10594, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10594
; LENGTH: 3019
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (115)..(144)
; OTHER INFORMATION: Band 7 protein family proteins domain identified by eMATRIX.
; OTHER INFORMATION: accession number BI01270C, p-value=6.745e-17, raw score of 16.91
; NAME/KEY: DOMAIN
; LOCATION: (2426)..(2903)
; OTHER INFORMATION: ATPases associated with various cellular act domain
; OTHER INFORMATION: Identified by Pfam, accession name AAA, E-value=3.3e-182, Pfam
; NAME/KEY: misc_feature
; LOCATION: (1)...(3019)
; OTHER INFORMATION: xaa - x or * as defined in Table 2
PCT-US01-08656-10594
```

Query Match 93.0%; Score 1643; DB 1; Length 3019;

Best Local Similarity 95.0%; Pred. No. 2e-141;
Matches 340; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

```
OY 1 MLARAARGHMGFPABG--LSTGFWPRSGRASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
```

DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118

OY 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178

DB 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178

OY 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238

DB 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238

OY 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEOYVSAFSKLAKDSNTILL 298

DB 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEOYVSAFSKLAKDSNTILL 298

OY 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOGTDASXDEIDRYKMT 356

DB 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOGTDASXDEIDRYKMT 356

RESULT 7

```
PCT-US01-04098A-1228
; Sequence 1228, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1228
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1228
```

Query Match 92.5%; Score 1635; DB 1; Length 358;

Best Local Similarity 94.7%; Pred. No. 3.9e-142;
Matches 341; Conservative 2; Mismatches 11; Indels 6; Gaps 3;

```
OY 1 MLARAARGHMGFPABG--LSTGFWPRSGRASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTVVLFPVPOEAVNVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLNDNTLQIDGVLRLINDPYKASYG 118
OY 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178
DB 119 VEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTIVDAINQAADCGIRCLREYEIK 178
OY 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238
DB 179 DIHVPREVESQMOMQVEARRKRAVLESEGTRESAINVAEGKKOQIILASEEKAEOIN 238
```


Db 179 DIHVPRKESQMOVFEARRRATVLESEGTRESAINVAEGKQAOILASAEKAEQIN 238
QY 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILAKDSNTIIL 298
Db 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILAKDSNTIIL 298
QY 299 PSNPGDVTSMAVAQAMGVYALTGA--PVPGTDSLSSGSSRDVGTGTDASXDEELDVRKMS 356
Db 299 PSNPGDVTSMAVAQAMGVYALTGA--PVPGTDSLSSGSSRDVGTGTDASXDEELDVRKMS 358

RESULT 8

PCT-US98-12125-272
; Sequence 272, Application PC/TUS9812125
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; NUMBER OF SEQUENCES: 318
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12125
; FILING DATE: Jan 01, 1900
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PZ008complete
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 272:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-272

Query Match 84.5%; Score 1493; DB 1; Length 306;

Best Local Similarity 99.3%; Pred. No. 4.3e-129;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 RMGRFHRIIEPGNLITLPIVLRIRYQSLKEIYINPEOSAVLDMVNTLOIDCVLRLIM 110
Db 1 RMGRFHRIIEPGNLITLPIVLRIRYQSLKEIYINPEOSAVLDMVNTLOIDCVLRLIM 60
QY 111 DPKASYGEDEYAVTQAOITMRSELGKLSXDKVFERESINASIYDAINCAADCGMI 170
Db 61 DPKASYGEDEYAVTQAOITMRSELGKLSXDKVFERESINASIYDAINCAADCGMI 120
QY 171 RCLRYEIKIHVPRKESQMOVFEARRRATVLESEGTRESAINVAEGKQAOILASE 230
Db 121 RCLRYEIKIHVPRKESQMOVFEARRRATVLESEGTRESAINVAEGKQAOILASE 180
QY 231 AEKAEQINQAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILA 290
Db 181 AEKAEQINQAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILA 240
QY 291 KDSNTILLPSNPGDVTSMAVAQAMGVYALTGA--PVPGTDSLSSGSSRDVGTGTDASXDEEL 350

Db 241 KDSNTILLPSNPGDVTSMAVAQAMGVYALTGA--PVPGTDSLSSGSSRDVGTGTDASXDEEL 300
QY 351 DRYKMS 356
Db 301 DRYKMS 306

RESULT 9

US-09-209-462B-411
; Sequence 411, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: PZ008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02

```
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-411

Query Match      84.5%: Score 1493; DB 16; Length 306;
Best Local Similarity 99.3%: Pred. No. 4,3e-129;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 51 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 110
DB 1 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 60
OY 111 DPKASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 170
DB 61 DPKASYGVDEPEYAVTQLAQTMRSELGKLSLDKYFRERESLNASTVDAINQADCMGI 120
OY 171 CLRREIKDHYPPRVKESOMQVEAERRRKATVLESEGTRESAIVNAEGKKOQILASE 230
DB 121 CLRREIKDHYPPRVKESOMQVEAERRRKATVLESEGTRESAIVNAEGKKOQILASE 180
OY 231 AEKAEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQVYSAFSKLA 290
DB 181 AEKAEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQVYSAFSKLA 240
OY 291 KDSNTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSSSSRDVQGTDAKDDEL 350
DB 241 KDSNTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSSSSRDVQGTDAKDDEL 300
OY 351 DRVKMS 356
DB 301 DRVKMS 306

RESULT 10
US-09-298-733-36
; Sequence 36, Application US/09298733
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffman, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapielko, Peter
; APPLICANT: Adams, R. Mark
; APPLICANT: Agostino, Michael J.
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steve H.
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding Them
; FILE REFERENCE: AG199-02PM
; CURRENT APPLICATION NUMBER: US/09/298,733
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/082,961
; EARLIER FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: 60/086,402
; EARLIER FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: 60/088,994
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/089,163
; EARLIER FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/091,619
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-298-733-36

Query Match      84.2%: Score 1488; DB 16; Length 305;
Best Local Similarity 99.3%: Pred. No. 1,2e-128;
Matches 303; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 52 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 111
DB 1 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 60
OY 112 PKKASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 171
DB 61 PKKASYGVDEPEYAVTQLAQTMRSELGKLSLDKYFRERESLNASTVDAINQADCMGI 120
OY 172 CLRREIKDHYPPRVKESOMQVEAERRRKATVLESEGTRESAIVNAEGKKOQILASE 231
DB 121 CLRREIKDHYPPRVKESOMQVEAERRRKATVLESEGTRESAIVNAEGKKOQILASE 180
OY 232 EKAQEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQVYSAFSKLA 291
DB 181 EKAQEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQVYSAFSKLA 240
OY 292 DSNFTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSSSSRDVQGTDAKDDEL 351
DB 241 DSNFTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSSSSRDVQGTDAKDDEL 300
OY 352 RVKMS 356
DB 301 RVKMS 305

RESULT 11
US-09-298-733A-36
; Sequence 36, Application US/09298733A
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapielko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding Them
; FILE REFERENCE: 1966.1009-000
; CURRENT APPLICATION NUMBER: US/09/298,733A
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-733A-36

Query Match      84.2%: Score 1488; DB 16; Length 305;
Best Local Similarity 99.3%: Pred. No. 1,2e-128;
Matches 303; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 52 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 111
DB 1 MGRFHRILPEGLNLIPLVDRIIRYVOSLKEIYINPEQSAVTLNDVNTLQIDGVLRLIM 60
```

Qy	112	PKKASYGEDEBEVAVTOLAQOTHTMSSEGLKLSXDKVFPRRESLMSIYDAINQADCMGIR	171
Db	61	PKKASYGEDEBEVAVTOLAQOTHTMSSEGLKLSLDVFPFRRESLMSIYDAINQADCMGIR	120
Qy	172	CLFREIKCIIHVPVRKESMQQVQAEERRKATVLESEGTRESAINVAEGKKQAQILNSEA	231
Db	121	CLFREIKCIIHVPVRKESMQQVQAEERRKATVLESEGTRESAINVAEGKKQAQILNSEA	180
Qy	232	EKAQCIQNAAGEASAVLAKAKAKAEAIRILLAALTOHNGDAASLVYAEQYVSASFSLAK	291
Db	181	EKAQCIQNAAGEASAVLAKAKAKAEAIRILLAALTOHNGDAASLVYAEQYVSASFSLAK	240
Qy	292	DSNTILLPNSNGDVTSMVAQAMGVYGAALTAKAPVGPDPDLSSSGSSRPVQGTDA\$XDEELD	351
Db	241	DSNTILLPNSNGDVTSMVAQAMGVYGAALTAKAPVGPDPDLSSSGSSRPVQGTDA\$XDEELD	300
Qy	352	RYKKS 356	
Db	301	RYKKS 305	

```

RESULT 12
US-09-723-594-36
; Sequence 36, Application US/09723594
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapleto, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding Them
; FILE REFERENCE: 1966.1009-002
; CURRENT APPLICATION NUMBER: US/09/723,594
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-594-36

```

	Query Match	84.2%	Score 1488	DB 21	Length 305
	Best Local Similarity	99.3%	Pred. 1.2e-128		
	Matches 303	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy	MGFRFRILEPGCINILIPVLDRIIRYQSLKEIYINPDEQSAVTLNDVTLIDGCVLYLRIMD	111			
Dd	1 MGRFRIRILEPGINILIPVLDRIIRYQSLKEIYINPDEQSAVTLNDVTLIDGCVLYLRIMD	60			
Qy	112 PYKASYGVEDPEYAVTOLAQTTRMSELGKLSXDKYFRERESINASIYDAINQAAQDCWGR	171			
Dd	61 PYKASYGVEDPEYAVTOLAQTTRMSELGKLSIDKYFRERESINASIYDAINQAAQDCWGR	120			
Qy	172 CLRREIKDHPVPRKESQMOVEARRRRKRAVYLSEGRRESAINVAEKKQAQIILASAA	231			
Dd	121 CLRREIKDHPVPRKESQMOVEARRRRKRAVYLSEGRRESAINVAEKKQAQIILASAA	180			

QY	223	EKAEDINQAGASAVLAKAKAKAEAIRLLAAALTOHNGDAAASLTVAEQVVSFSLAK	231
Db	181	EKAEDINQAGASAVLAKAKAKAEAIRLLAAALTOHNGDAAASLTVAEQVVSFSLAK	240
QY	292	DSNTLLPSNPEDVYSMTAOAMGYGALTTPKAPVGTGPDSTSSGSSRPVQGTDAAXDEELD	355
Db	241	DSNTLLPSNPEDVYSMTAOAMGYGALTTPKAPVGTGPDSTSSGSSRPVQGTDAAXDEELD	300
QY	352	RYKMS	356
Db	301	RYKMS	305

```

RESULT 13
US-09-724-497-36
; Sequence 36, Application US/09724497
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapijko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE OF INVENTION: Encoding Them
; FILE REFERENCE: 1966.1009-001
; CURRENT APPLICATION NUMBER: US/09/724,497
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-497-36

```

Query Match	84.2%	Score 1488	DB 21	Length 305
Best Local Similarity	99.3%	Pred. No. 1,2e-128		
Matches 303	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	52	MGRFHRILEPGINILIPVLDRIYVQSLKEIYINNEQSAVTLNDVTLIDGVLVLRIMD	111	
Db	1	MGRFHRILEPGINILIPVLDRIYVQSLKEIYINNEQSAVTLNDVTLIDGVLVLRIMD	60	
QY	112	PKYASGYEDPEXAVYQLAQTMRSELGKLSXDKYFRERESINASTVDAINQAADCGIR	171	
Db	61	PKYASGYEDPEXAVYQLAQTMRSELGKLSXDKYFRERESINASTVDAINQAADCGIR	120	
QY	172	CRYEIKDIIHPPRKESQMOMVEAEERRKRAVLSEGRRESAINVAEKKKQAOIILASEA	231	
Db	121	CRYEIKDIIHPPRKESQMOMVEAEERRKRAVLSEGRRESAINVAEKKKQAOIILASEA	180	
QY	232	EKAEOINQAGEASVILAKAKAEAIRLLAALTOHNDAAASTLVAAOYSAFSLAK	291	
Db	181	EKAEOINQAGEASVILAKAKAEAIRLLAALTOHNDAAASTLVAAOYSAFSLAK	240	
QY	292	DSNTTILLPSPNDVYTMVAQAMGVYCALTKAPVPGPDSLSGSSSRDVOGTDA SXDELD	351	
Db	241	DSNTTILLPSPNDVYTMVAQAMGVYCALTKAPVPGPDSLSGSSSRDVOGTDA SXDELD	300	
QY	352	RVKMS	356	

Db 301 RVKMS 305

RESULT 14
US-09-614-150-5415

```
; Sequence 5415, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/09/614,150
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5415
; LENGTH: 323
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-09-614-150-5415
```

Query Match 60.2%; Score 1064.5; DB 20; Length 323;
Best Local Similarity 63.7%; Pred. No. 2,1e-89;
Matches 209; Conservative 55; Mismatches 47; Indels 17; Gaps 2;

```
QY 39 LFVPOQEAAMVVERMGFRHRLLEPGNLILIPVLRIRYQSLKEIYINPEQSAVTLDNVT 98
   |||
   1 MFVPOQEAAMVVERMGFRHRLLEPGNLILIPVLRIRYQSLKEIYINPEQSAVTLDNVT 60
QY 99 LQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTV 158
   |||
   61 LSIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTV 120
QY 159 DAINQAOCWGRCLRYELKIDIHVPRKESQMOVEAEERRKRAVLSEGTRESAIVNA 218
   |||
   121 DSINKASAMGJACLRYEIRIDRLPTRVHEAMQOMOVEAEERRKRAVLSEGTRESAIVNA 180
QY 219 EGKROAQLIASEAEKAEQIINQAGEASAVLAKAKAKAIRILAAALTOHNGDAASLTV 278
   |||
   181 EGKRSRLILASEAEKAEQIINQAGEASAVLAKAKAKAIRILAAALTOHNGDAASLTV 240
Db 279 AEQYVSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVYGALTKAPVPGTDSLSSGSSRD 338
   |||
   241 AEQYIGAFKKLAKTNTMTILPSNPGDVNGFVAQALAVYNHVSNS-----NQATKSSSEN 293
QY 339 VQGTDA-----SXDELDRLRYKMS 356
   |||
   294 VKGVGACLNASVEYKELQEDKSSVKMN 321
Db
```

RESULT 15
US-60-191-637-5433
; Sequence 5433, Application US/60191637

```
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5433
; LENGTH: 323
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-60-191-637-5433
```

Query Match 60.2%; Score 1064.5; DB 26; Length 323;
Best Local Similarity 63.7%; Pred. No. 2,1e-89;
Matches 209; Conservative 55; Mismatches 47; Indels 17; Gaps 2;

```
QY 39 LFVPOQEAAMVVERMGFRHRLLEPGNLILIPVLRIRYQSLKEIYINPEQSAVTLDNVT 98
   |||
   1 MFVPOQEAAMVVERMGFRHRLLEPGNLILIPVLRIRYQSLKEIYINPEQSAVTLDNVT 60
QY 99 LQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTV 158
   |||
   61 LSIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTMRSELGKLSXDKVFRERESLNASTV 120
QY 159 DAINQAOCWGRCLRYELKIDIHVPRKESQMOVEAEERRKRAVLSEGTRESAIVNA 218
   |||
   121 DSINKASAMGJACLRYEIRIDRLPTRVHEAMQOMOVEAEERRKRAVLSEGTRESAIVNA 180
QY 219 EGKROAQLIASEAEKAEQIINQAGEASAVLAKAKAKAIRILAAALTOHNGDAASLTV 278
   |||
   181 EGKRSRLILASEAEKAEQIINQAGEASAVLAKAKAKAIRILAAALTOHNGDAASLTV 240
Db 279 AEQYVSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVYGALTKAPVPGTDSLSSGSSRD 338
   |||
   241 AEQYIGAFKKLAKTNTMTILPSNPGDVNGFVAQALAVYNHVSNS-----NQATKSSSEN 293
QY 339 VQGTDA-----SXDELDRLRYKMS 356
   |||
   294 VKGVGACLNASVEYKELQEDKSSVKMN 321
Db
```

Search completed: September 22, 2002, 18:26:25
Job time: 728 sec

